MODELING, ANALYSIS, AND OPTIMIZATION OF CACHE-ASSISTED WIRELESS NETWORKS- A STOCHASTIC GEOMETRY APPROACH

BY

MOSTAFA MOHAMED LOTFY EMARA

A Dissertation Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In

ELECTRICAL ENGINEERING

MAY 2018

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by MOSTAFA MOHAMED LOTFY EMARA under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING**.

Dissertation Committee

Dr. Samir Al-Ghadban (Adviser)

Dr. Tareq Al-Naffouri (Member)

Dr. Azzedine Zerguine (Member)

Dr. Ali H Muqaibel (Member)

Dr. Ashraf Sharif Mahmoud (Member)

Dr. Ali A. Al-Shaikhi
Department Chairman

Dr. Salam A. Zummo
Dean of Graduate Studies

23/7/2018
Date

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

To the memory of my father
my mother, brother, and sisters
my lovely wife, Zaid, and Maria

ACKNOWLEDGMENTS

Firstly, Praise be to Allah, Lord of the Worlds, for the good health and wellbeing that were necessary to complete this dissertation.

I would like to express my sincere gratitude to my advisor Dr. Samir Al-Ghadban for his close supervision and his continuous support along my thesis journey.

Special thanks to my committee members of Dr. Tareq Al-Naffouri, Prof. Azzedine Zerguine, Dr. Ali H Muqaibel, and Dr. Ashraf Mahmoud for their constructive comments that surely improve the quality of thesis.

I would like to express my deeply gratitude to Dr. Tareq Al-Naffouri, Dr. Hesham ElSawy, and Dr. Sameh Sorour for their guidance, patience, motivations and providing me with the key concepts that help me complete this work.

I would also like to thank my mother, brother, sisters, and my friend Dr. Hossam Younes for always supporting and encouraging me with their best wishes.

Special thank for my beloved wife and my children Zaid and Maria who suffer alot during this journey.

Finally, I would like to express my thanks to my Electrical Engineering Department and my KFUPM University for their friendly support and cooperation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

SBS	:	Small Base Station
KPI	:	Key Performance Indicator
OSA	:	Opportunistic Spectrum Access
5G	:	Fifth Generation
D2D	:	Device-to-Device
SINR	:	Signal-to-Interference-Plus-Noise-Ratio
SIR	:	Signal-to-Interference Ratio
SNR	:	Signal-to-Noise-Ratio
mmW	:	Millimeter-Wave
PPP	:	Poisson Point Process
LT	:	Laplace Transform
MBS	:	Macro Base Station
DASH	:	Dynamic Adaptive Streaming over HTTP
VoD	:	Video on Demand
eMBMS	:	Evolved Multimedia Broadcast and Multicast Services
FDM	:	Frequency Division Multiplexing

PDF	:	Probability Density Function
PMF	:	Probability Mass Function
NCL	:	Neighboring Cell List
LOS	:	Line of Sight
NLOS	:	Non-Line of Sight
CCDF	:	Counterpart Cumulative Density Function
KKT	:	Karush- Kuhn-Tucker
SDM	:	Spatial Division Multiplexing
CAPEX	:	Capital Expenditure
mmFN	:	mmW Fog Network
CE-BS	:	Cache Enabled Base Station
CF-BSs	:	Cache-Free Base Station

THESIS ABSTRACT

NAME:	Mostafa Mohamed Lotfy Emara
TITLE OF STUDY:	Modeling, Analysis, and Optimization of Cache-Assisted Wireless Networks- A Stochastic Geometry Approach
MAJOR FIELD:	Electrical Engineering
DATE OF DEGREE:	May 2018

Cache-enabled small base station (SBS) densification is foreseen as a key component of 5G cellular networks. This architecture enables storing popular files at the network edge (i.e., SBS caches), which empowers local communication and alleviates traffic congestions at the core/backhaul network. The hit probability, which is the probability of successfully transmitting popular files request from the network edge, is a fundamental key performance indicator (KPI) for the cache-enabled network. Maximizing the hit probability requires both an optimal content placement in the SBSs and an efficient spectrum/interference management. However, the literature mostly focuses on the optimal caching problem and overlooks the appropriate spectrum and interference management aspects of cellular networks.

We propose the use of opportunistic spectrum access (OSA) in caching systems, which was not previously considered in the literature. Exploiting the available multi-channels that are universally reused by all SBSs, the catering SBS opportunistically transmits the requested file on a channel that is not used by any closer SBS to the requesting user. This would increase the hit probability in case the requested file is not cached in a closer SBS.

We develop a scheduling aware mathematical framework, based on stochastic geometry, to characterize the hit probability of single-user cache-enabled network in a multi-channel environment with OSA. To this end, we assess and compare the performance of two widely employed caching distribution schemes, namely, uniform caching and Zipf caching. The numerical results show that the commonly used single channel environment leads to pessimistic assessment for the hit probability of cache-enabled network. Furthermore, the numerical results manifest the superiority of the Zipf caching scheme and quantify the hit probability gains in terms of the number of channels and cache size.

We model and analyze a multi-channel cache-enabled 5G networks with both unicast/multicast capabilities and OSA. To this end, we first derive the hit probability. An optimization framework for file caching is then developed to maximize this hit probability. To this end, a simple concave approximation for the hit probability is proposed, which highly reduces the optimization complexity and leads to a closed-form solution. The sub-optimal solution is benchmarked against the

uniform and Zipf caching, through numerical results and extensive simulations. It is shown that the caching strategy should be adapted to the network parameters and capabilities. For instance, diversifying file caching according to the Zipf distribution is better in multicast systems with large number of channels. However, when the number of channels is low and/or the network is restricted to unicast transmissions, it is better to confine the caching to the most popular files only.

Finally, we propose a flexible caching scheme for the mmWave partially cache-enabled network in which the cache size is inversely proportional to the intensity of cache-enabled SBSs. The optimal caching parameters (i.e., caching distribution and the cache-enabled fraction) that maximizes the hit probability is obtained. To this end, the optimal solution is benchmarked against the most popular caching. The trade-off between the cache-size and the intensity of cache-enabled SBSs is addressed via simulation results. The optimal caching outperforms the most popular caching thanks to both the dependency of the network parameters and the files diversity.

ملخص الرسالة

الاسم الكامل: مصطفى محمد لطفى أحمد عمارة

عنوان الرسالة: نمذجة وتحليل وتعظيم أداء الشبكات اللاسكية

التخصص: الهندسة الكهربائية

تاريخ الدرجة العلمية: مايو 2018

تعد الشبكات المعتمدة على ذاكرة التخزين احدى المكونات الاساسية للجيل الخامس للشبكات الخلوية. تتيح هذه الشبكات تخزين الملفات عالية الطلب من مستخدمى الخدمة بذاكرة التخزين لدى وحدات الارسال الصغيرة المنتشرة . تتولى هذه الوحدات القريبة من المستخدمين خدمتهم مباشرة بدلا من خدمتهم عبر الخادمات المركزية البعيدة وهذا من شأنه يعزز الخدمة ويقلل الضغط على كابلات التواصل مع الخادمات المركزية. تعد احتمالية خدمة المستخدم عبر هذه الوحدات هو الآلية للحكم على مثل هذه الشبكات. علما بأنه يتم خدمة المستخدم من خلال اقرب وحدة تحتوى الملف المطلوب فى ذاكرتها والتى من الممكن الا تكون اقرب وحدة منه مكانيا. تبنى هذا الأسلوب من شأنه يزيد التداخلات وخاصة من الوحدات الاقرب للمستخدم من وحدته الخادمة.

أولا: فى هذه الرسالة نسعى التعظيم هذه الاحتمالية من اجل تحسين أداء الشبكة ولذلك تم استخدام تقنية خدمة المستخدم على تردد مختلف عن المستخدم من قِبل الوحدات الأقرب منه مكانيا. تبنى هذه التقنية تؤدى الى تقليل التداخلات عند المستخدم وبالتالى التحسين العام لمستوى الخدمة.

ثانيا: دراسة كيفية توزيع الملفات داخل وحدات الارسال معتمدة فى ذلك على معدل طلب المستخدمين لكل من هذه الملفات. دمج آلية الارسال على ترددات مختلفة مع التوزيع الأمثل للملفات داخل الوحدات من شأنه تعظيم احتمالية خدمة المستخدم عبر هذه الوحدات دون الحاجة لخدمته عبر الخادمات المركزية.

ثالثا: تم حل مسألة تعظيم لإيجاد الآلية المثلى لتوزيع الملفات داخل الوحدات وتم مقارنتها بالآليات التقليدية المتعارف عليها وأظهرت النتائج افضلية للنظام المقترح ولا سيما فى ظل تجنب التداخلات بتبنى الآلية المذكورة آنفاً.

رابعا: تم دراسة الشبكة المعتمدة على ذاكرة التخزين والمستخدمة لتقنية ارسال الاشارات عالية الترددات (millimeter wave) وتم تبنى مقترح من شأنه ربط متوسط الوحدات ذوى ذاكرة التخزين بسعة ذاكرة التخزين لديهم. فكلما زاد هذا المتوسط كلما قلت سعة التخزين. تم ايضا ايجاد التوزيع الامثل للملفات الى جانب النسبة المثلى للوحدات ذوى ذاكرة التخزين. تم مقارنة الآلية المقترحة بالآليات المتواجدة لإظهار تفوق المقترح.

CHAPTER 1

INTRODUCTION

1.1 Motivation

The rapid proliferation of social networking along with the advancement in smart devices increase the traffic burden on cellular operators. Such evolution entails massive numbers of connections with data hungry application. For instance, video streaming is a main contributor to the traffic burden, with anticipations to consume around 82% of the total internet traffic by 2020 [2]. To address that, the fifth generation (5G) of cellular networks dictates tangible performance leap in terms of network capacity (100 fold) and transmission rate (1000 fold) [3]. Such performance requirement is expected to be fulfilled by an unprecedented network densification phase, in which the SBS density can reach hundreds of SBSs per km^2 [4], [5]. Deploying more SBSs in the same geographical location improves the spatial frequency reuse and reduces the number of users associated to the same BS, thus fostering light loaded BSs with high data rate links. However, supplying

this large number of SBSs with real-time data services impose a huge burden on the backhaul links. Consequently, backhauling represents a non-trivial bottleneck to attain the foreseen network densification gains.

Fortunately, with the aid of data analytics on social networks, operators can estimate the popularity of contents (e.g., videos, mobile-applications, and software updates) in terms of the volume of user demand. Relying on this fact, the 5G of cellular networks is evolving towards data-centric networking [6–9]. In particular, cellular operators will densify their networks via cache-enabled small base stations (SBSs), and implement proactive caching of the popular content at the network edge [10–12]. Furthermore, proactive caching at the device-to-device (d2d) level is also proposed to fully utilize network edge resources [13–16]. This clearly improves the spatial frequency reuse, relieves the backhaul congestion, and reduces download latency. For instance, recent studies show that proactive caching has the potential to reduce 66% of the backhaul traffic when employed in cellular networks [17]. Fig. 1.1 depicts the concept of caching at the network edge. The dashed lines represent the local transmission links, while the bold red line represents the backhaul link. Fig. 1.1 shows that the requested file is either downloaded from the local cache or through the core network over the backhaul link depending on whether the serving SBS stores the file or not.

A fundamental and widely employed design objective in data-centric networks is to maximize the probability of users finding requested popular files in the cache of an accessible edge SBS, also known as the *hit probability*. The high density of

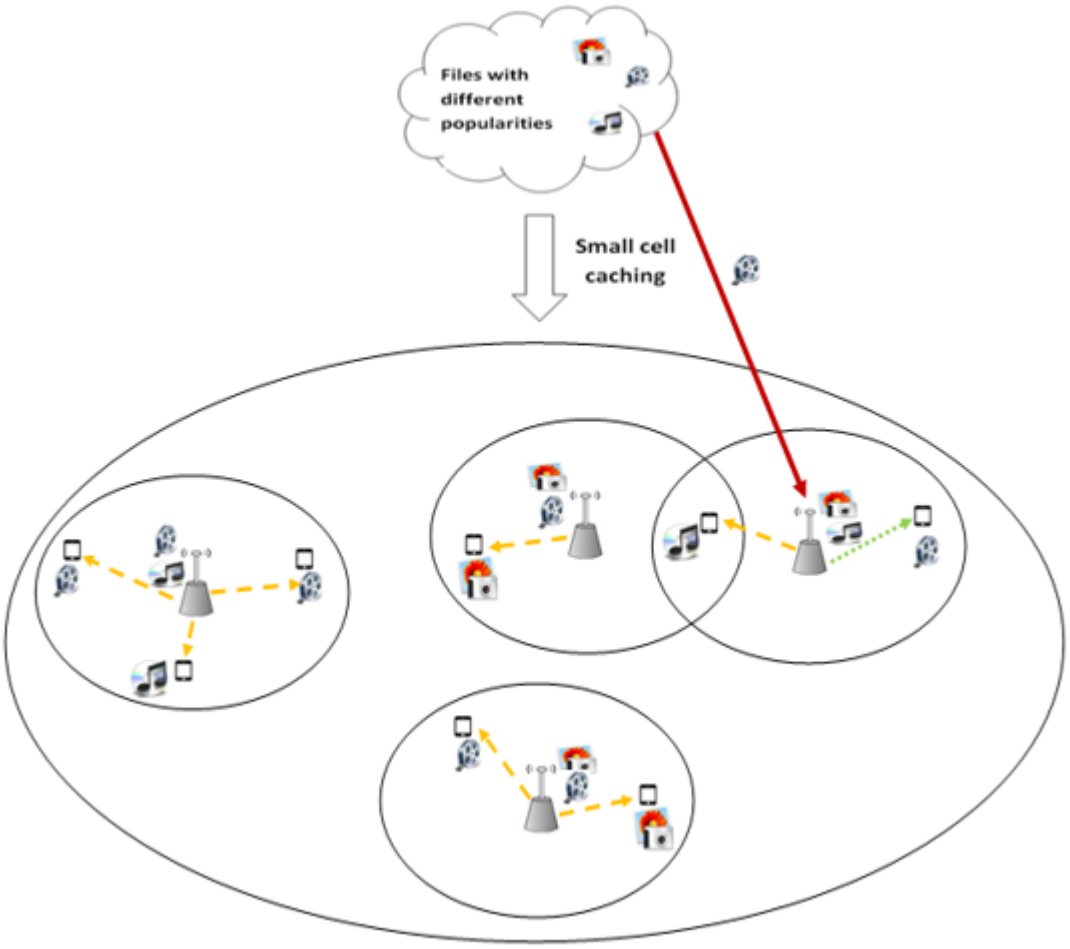


Figure 1.1: An illustration of cache-enabled network

SBSs diversifies the association opportunities of users. Hence, the hit event is not confined to the geographically closest SBS. More precisely, the hit probability is defined by the joint event of finding a SBS that i) is caching the requested popular file; and ii) is capable of serving the requesting user with at least a minimum pre-defined threshold of signal-to-interference-plus-noise-ratio (SINR). Such joint event implies that maximizing the hit probability requires both an optimal content placement in the SBSs and an efficient spectrum/interference management. On one hand, optimal caching increases the probability to find the requested content

in a proximate SBS. On the other hand, since the catering SBS is not necessarily the geographically closest SBS to the requesting user, spectrum and interference management is necessary to attain the predefined SINR threshold. However, the literature mostly focuses on the optimal caching problem and overlooks the appropriate spectrum and interference management aspects of cellular networks. Consequently, in chapters 2 and 3 we adopt the OSA strategy to mitigate strong interference received at the user from the SBSs geographically closer to it than its serving SBS. Moreover, the optimal caching that maximizes the hit probability is derived for both the *unicast* and *multicast* transmission schemes.

1.2 Millimeter-Wave Communication

Another agile approach that was proposed in the literature to address the high capacity requirements is to exploit the unused or partially used millimeter-wave (mmWave) bands (20-100 GHz). However, such frequencies expose to large path loss and poor penetration (blockage) through concrete, foliage, and other common materials that primarily confines the mmWave based cellular communication. Recently, research efforts show that the mmWave transmission within ranges of 150-200m in a dense urban environment is feasible [18–20]. Thanks to the high gain antenna arrays that are capable to be equipped at the BSs and end-user terminals due to the small wavelengths. This compensates the increased path loss and blockage effects, overcomes the additional noise due to the large bandwidth, and isolates the other-cell interference. Exploiting the narrow-beam transmissions, the

selfbackhauling concept was introduced in [21] in which the BSs that have wired backhaul links with the core network provide wireless backhaul to those have not wired backhaul links.

Recently, cache-enabled mmWave cellular networks have been studied in the literature [22–24] in which cache-enabled SBSs are densely deployed to overcome the high path-loss and blockage effects. However, supplying all the SBSs in the network with storage capacity enlarges the capital expenditure. Otherwise, SBSs are equipped with small-size caches, which highly limits the diversity in caching distribution and thus leads to a pessimistic performance assessment. From another perspective, partially cache-enabled network is a likely deployment scenario in the transition to a fully cache-enabled cellular network. Motivated by the aforementioned, a flexible caching strategy for a partially cache-enabled mmWave network is proposed in chapter 4. According to this strategy, the caches size is inversely proportional to the intensity of the cache-enabled SBSs. In addition, we exploit the *self backhauing* concept, which is introduced in [21], to provide the serving SBS with the user's requested popular content.

1.3 Stochastic Geometry

In this section, we introduce a brief review on stochastic geometry, the mathematical tool that will be utilized throughout this work to characterize the system performance. Stochastic geometry is a widely used tool for modeling and analyzing cellular networks [25]. In stochastic geometry analysis, a practical and

tractable point process is chosen to abstract the network. In the literature, the Poisson point process (PPP) is widely accepted for modeling cellular networks due to its simplicity and tractability [26, Table I] and has been verified in [27–29] by several empirical studies. In stochastic geometry analysis, it is very important to characterize the service distance distribution, the Laplace Transform (LT) for the aggregate interference, and the SINR, which are further used to obtain other performance metrics like coverage probability and ergodic capacity. The hit probability, the performance metric for the cache-enabled network, depends on the coverage probability as it will be explained further in details. Here, we define some of the performance metrics, which are extensively studied in the stochastic geometry literature.

Coverage probability: It is defined as the probability that the received SINR is greater than some threshold (T) i.e., $P[SINR > T]$.

Spectral efficiency: It is the maximum rate per unit bandwidth that can be reliably transmitted over a channel. It is defined by Shannon's capacity formula given by: $C = log(1 + SINR)$.

Ergodic capacity: It is the long term average achievable rate averaged over all channel states. It can be expressed as $E[log(1 + SINR)]$.

The rest of this chapter includes the literature survey in section 1.4, the thesis contributions and the outcome publications are presented in section 1.5 and 1.6, respectively. Finally, the thesis organization is introduced in section 1.7.

1.4 Literature Survey

Caching at the network edge has been introduced by Golrezaei *et al.* in their papers [10,30–32]. The authors focus on the content placement problem and they discussed several ways of solving the problem under the assumption of coded and uncoded caching of files. The objective was to find the optimum distribution of files (or part of files) among the cache-enabled SBSs such that the amount of date rate served by the macro base station (MBS) is minimized. The proposed scheme were examined under several environments and technologies, and throughput enhancements were shown by simulations. In [33], a joint caching algorithm is developed for a two-layer network exploiting both the single-layer multicast opportunities and the correlations of caching contents across the two layers. However, the works in [10, 30–33] assume interference free environment. Using stochastic geometry, the interference effect is accounted for in [34, 35], where the policy of caching most popular files is adopted. Thus, the user is always associated to its geographical closest SBS and the requested file is either downloaded from the local cache or through the core network over the backhaul link. However, such model lacks the caching diversity, underutilizes the dense cellular infrastructure, and does not exploit multiple SBS diversity in terms of association. In [1, 36–39], the optimal file placement in a single-tier cache-enabled network is proposed to maximize the hit probability. In [40, 41], a multi-tier network consists of cache-enabled users, relays (SBSs), and MBSs tiers is analyzed and expressions for the average ergodic rate, throughput, and the delay are derived. Due to the prob-

abilistic caching distributions in [1, 36–41], the user is associated to the closest SBS that stores its requested file. However, the works in [1, 36–41] consider a single channel system in a fully loaded SBS scenario, which contradicts with the intrinsic multi-channel nature of cellular networks and thus leads to pessimistic performance assessments caused by the strong interference received at the user from the SBSs closer to it than its serving SBS.

In practice, the multi-channel access in a cellular network is exploited to diversify transmissions, relief interference, and thus, improve the network performance. The multi-channel exploitation is attained through spectrum access strategies that guarantee efficient usage of the spectrum resources. In this context, OSA was shown to be an appealing solution to mitigate inter-cell interference and boost the coverage probability [42–45]. It worth noting that having more than one radio channel at the SBSs will not alleviate the need for popular file caching at the SBSs to reduce the backhaul congestion. On the contrary, combining OSA with caching at the SBSs can further reliefs such congestion by providing more acceptable transmission rates and SBS coverage (and thus access to more cached files) to the user. Indeed, the scenario of caching with universal frequency reuse, high loads, and no OSA, highly limits the transmission rates due to operating at very-low SINRs (strictly less than 0 dB). This in turns limits the coverage of each SBS, reduces the hit probability (as each user will have access to less SBSs), and thus adds more burden on the backhaul links to download the requested files from the core networks to the closest SBSs of their requesting users. Depending

on the application and network capabilities, two approaches can be employed to deliver the popular content, namely, the *unicast* and *multicast* schemes. The unicast scheme serves each user request on a unique frequency channel irrespective of the requested content. In contrast, multicast schemes group similar requests and serve them all on the same channel. Hence, multicast schemes are more spectrum efficient, but at the expense of coordination and synchronization overhead. Due to the imposed tradeoff between complexity and spectral efficiency, both the unicast and multicast schemes are used in practice. For instance, the dynamic adaptive streaming over HTTP (DASH) is a unicast video delivery technique that is used for video on demand (VoD) services [46]. On the other hand, the Evolved Multimedia Broadcast and Multicast Services (eMBMS) technology is utilized within the 3GPP LTE for Live Video/Audio streaming [47]. Hybrid multicast/unicast transmission techniques also exist for DASH over MBMS for LTE [48]. Since both techniques are viable alternatives for popular content delivery in 5G networks, they are both studied in the literature. For instance, a joint optimal file placement and multicast in a multi-channel system is proposed in [49–51] for single and multi tiers network. However, they follow a frequency division multiplexing (FDM) scheme in which each SBS divides the shared available bandwidth among its tagged users. Such rigid frequency reuse scheme leads to strong interference received at the user from the closest SBSs closer to it than its serving SBS, which is well-known to underutilize the spectrum resources [45].

In the literature, the *uniform caching* (i.e., the popular files are distributed into

the SBSs randomly irrespective to their popularity.), *popularity-based caching* (i.e., the files are stored into the caches according to their popularity), and the *most popular caching* (i.e., the SBSs cache the M most popular files, where M is the cache size) are the widely known schemes to distribute the files into SBSs caches. They are commonly considered as the benchmarks file placement schemes [49–51]. The work in [52] shows that the popularity-based caching outperforms the uniform caching in terms of the hot probability for a given minimum cache size. In [53], the authors show that the optimal caching distribution for a d2d cache-enabled network follows the same trend as the files popularity distribution. More precisely, the optimal caching for a Zipf-distributed files popularity can also be modeled as a Zipf distribution with exponent depends on both the files popularity exponent and the network parameters.

Several stochastic geometry based studies to model and analyze the mmWave cellular network are existed in the literature [21, 54–56]. In [54], the authors analyze a single-tier mmWave-based network to characterize the rate coverage probability in which a single-ball approximation model for the blockages is proposed. In [55,56], the coverage probability for multi-tier mmWave cellular network is analyzed. Inspired by the experimental measurements in [57], the single-ball blockage approximation model in [54] is upgraded to a two-ball and a generalized d-ball models in [55] and [56], respectively. In [55, 56], the impact of the beamforming misalignment on the coverage probability is investigated to get insight on the performance in practical scenarios. They also addressed the performance

of a hybrid cellular network operating in both mmWave and sub-6GHz frequency bands. However, [21, 54–56] adopt a flat-top antenna gain pattern to maintain analytical tractability. In [58], the coverage probability for the mmWave cellular network is derived considering more sophisticated antenna gain patterns to approximate the actual antenna pattern.

Motivated by the fact that caching at the network edge is one of 5G service requirements and mmWave will be a key component of future wireless access [59], cache-enabled mmWave cellular network was proposed in the literature [22–24,60]. In [61], caching distribution for video streaming contents into SBSs' caches on the highway is studied and it is shown to significantly reduce the connection and retrieval delays. The optimal caching distribution that maximizes the successful content delivery probability is derived for a hybrid mmWave and sub-6 GHz cache-enabled network in [22, 24]. The authors in [23] infer that the most popular caching policy is the optimum to maximize the area spectral efficiency for a hybrid system of traditional backhaul-connected SBSs and cache-enabled SBSs.

1.5 Thesis Contributions

Based on the aforementioned discussions, the thesis contributions can be summarized as follows:

- We propose the use of OSA in caching systems, which was not previously considered in the literature. Exploiting the available multi-channels that are universally reused by all SBSs, the catering SBS opportunistically transmits

the requested file on a channel that is not used by any closer SBS to the requesting user. This would increase the hit probability in case the requested file is not cached in a closer SBS.

- We develop a scheduling aware mathematical framework, based on stochastic geometry, to characterize the hit probability of single-user cache-enabled network (i.e., hybrid services network) in a multi-channel environment with OSA. To this end, we assess and compare the performance of two caching schemes, namely, the uniform and Zipf caching. Closed-form expressions are obtained for the coverage and hit probabilities for a multi-channel caching scenario.

- We extend the work to a full multi-channel cache-enabled 5G networks with both unicast/multicast capabilities and OSA. An optimization framework for file caching is then developed to maximize this hit probability. To this end, a simple concave approximation for the hit probability is proposed, which highly reduces the optimization complexity and leads to a closed-form solution. The sub-optimal solution is benchmarked against two widely employed caching distribution schemes, namely uniform and Zipf caching, through numerical results and extensive simulations. It is shown that the caching strategy should be adapted to the network parameters and capabilities.

- We propose a flexible caching scheme for the mmWave partially cache-enabled network in which the cache size is inversely proportional to the

intensity of cache-enabled SBSs. Exploiting the self backhauling concept, which is introduced in the literature, the cache-enabled SBSs provide wireless bakhaul to those have not caches with their stored files. The optimal caching parameters (i.e., caching distribution and the cache-enabled fraction) that maximizes the hit probability is obtained. To this end, the optimal solution is benchmarked against the most popular caching.

1.6 Publications

Our contributions have led to the following publications:

- M. Emara, H. ElSawy, S. Sorour, S. Al-Ghadhban, M. S. Alouini and T. Y. Al-Naffouri, "Optimal Caching in 5G Networks with Opportunistic Spectrum Access," IEEE Transactions on Wireless Communications.

- M. Emara, H. ElSawy, S. Sorour, S. Al-Ghadhban, M. S. Alouini and T. Y. Al-Naffouri, "Optimal Caching in Multicast 5G Networks with Opportunistic Spectrum Access," IEEE Globecom 2017, Singapore, Dec 2017.

- M. Emara, H. ElSawy, S. Sorour, S. Al-Ghadhban, M. S. Alouini and T. Y. Al-Naffouri, "Stochastic geometry model for multi-channel fog radio access networks," 2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Paris, 2017, pp. 1-6.

- M. Emara, H. ElSawy, S. Sorour, S. Al-Ghadhban, M. S. Alouini and T. Y.

Al-Naffouri, "Flexible Design of Millimeter-Wave Cache Enabled Fog Networks," ***submitted to*** IEEE Globecom workshops 2018, Abu Dhabi, Dec 2018.

1.7 Thesis Organization

The thesis is organized as follows:

Chapter 1 introduces the motivations for this work, the literature survey of the related work, and the achieved contributions.

Chapter 2 analyzes and models the single-user cache-enabled network in a multi-channel environment with OSA. A mathematical paradigm is developed for the hit probability, which is the considered system performance metric in this thesis. The performance of two most popular caching schemes, namely, the uniform and Zipf caching is studied through independent Monte Carlo simulations.

Chapter 3 analyzes and models the multi-channel cache-enabled 5G networks with both unicast/multicast capabilities and OSA. The optimal caching distribution that maximizes the hit probability is obtained for both unicast/multicast transmission schemes. The sub-optimal solution is benchmarked against the uniform and Zipf caching distributions. The design insights for the developed paradigm are revealed.

Chapter 4 addresses the tradeoff between the cache size and the intensity of cache-enabled SBSs in a mmWave partially cache-assisted SBSs. The optimal caching parameters (i.e., the caching distribution and the cache-enabled fraction)

that maximizes the hit probability are derived. The optimal solution is benchmarked against the uniform and Zipf caching schemes.

Chapter 5 draws conclusions from the present work and discusses potential directions for future work

CHAPTER 2

ANALYSIS OF SINGLE-USER CACHE-ENABLED NETWORK WITH OPPORTUNISTIC SPECTRUM ACCESS

2.1 Introduction

This chapter develops a scheduling aware mathematical framework, based on stochastic geometry, to characterize the hit probability of cache-assisted network in a multi-channel environment. To this end, we assess and compare the performance of two caching schemes, namely, the uniform and Zipf caching. Closed-form expressions are obtained for the coverage and hit probabilities for a multi-channel caching scenario. These expressions generalize the single channel scenario that is

widely adopted in the literature. All theoretical results in this chapter are verified via independent Monte Carlo simulations. The numerical results manifest the pessimistic results of the single channel cache-enabled network and quantify the hit probability gains in terms of the number of channels, cache size, and caching distribution.

The rest of this chapter is organized as follows: Section 2.2 introduces the proposed system model. In Section 2.3, the OSA success probability is derived. The coverage probabilities when the OSA succeeds or fails are obtained in Section 2.4. Finally, the developed mathematical performance metric is verified via the numerical results in Section 2.5.

2.2 System Model

2.2.1 Network Model

We assume that the SBSs and the users are distributed in $\mathbb{R}^2$ according to two independent homogeneous PPPs, $\mathbf{\Phi}_b$ and $\mathbf{\Phi}_u$, with intensities λ_b and λ_u, respectively. In the considered downlink transmissions, all SBSs transmit with the same power P and share a common set of channels $\mathbf{S}$ with cardinality $|\mathbf{S}|$. The signal attenuation due to propagation is characterized by the power-law distance dependent path-loss model $r^{-\eta}$, where r is the propagation distance and $\eta > 2$ is an environment dependent exponent. A Rayleigh fading environment is assumed with independent and identically distributed channel gains with unit mean pow-

ers. According to Slivnyak-Mecke theorem [62], there is no loss of generality in focusing on a test user located at the origin, to analyze the network performance.

2.2.2 Caching Model

We consider a finite library of popular files (contents), denoted by $\mathbf{J} = \{c_1, c_2, ..., c_J\}$. It is assumed that all files have the same length. However, this analysis may still be applied with files of different sizes by chopping each file into equal length packets. We assume that the files popularity is fully known a priori for the network operator and follows the Zipf distribution due to its practical relevance [63]. The assumption of knowing the files popularity a priori is commonly used in the literature [1,36–39,49,50]. Such assumption is consistent with the fact that the rate at which the file popularity changes is most likely much lower than the rate they are requested with, otherwise there is no notion of popularity. Addressing estimation errors and time-varying characteristics of the file popularity is beyond the scope of this work. Therefore, the files popularity can be expressed as:

$$a_j = \frac{j^{-\gamma}}{\sum_{i=1}^{J} i^{-\gamma}} \tag{2.1}$$

where a_j is the probability that a generic user requests file c_j, and γ is the Zipf parameter that governs the popularity distribution skewness. Larger (smaller) γ increases (decreases) the discrepancies among the files popularity and implies that fewer (more) files are frequently requested. Without loss of generality, it is assumed that the files of the library are enumerated in a descending order of their

popularity, i.e., $a_1 \geq a_2 \geq \cdots \geq a_J$. Each SBS has a cache memory of size $M < J$ files. It is assumed that each SBS independently chooses a combination $\mathbf{x} \in \mathbf{X}$ of M different files to store in its cache, where $\mathbf{X} = \{1, 2, \ldots, |\mathbf{X}|\}$ denotes the set of all possible combinations with a set cardinality $|\mathbf{X}| = \binom{J}{M}$. File redundancy at the same cache is avoided as it wastes the memory resources at no additional benefit. Let $p_{\mathbf{x}}$ denote the probability that a generic SBS stores a combination $\mathbf{x} \in \mathbf{X}$. Consequently, $p_{\mathbf{x}}$ satisfies the following constraints:

$$\begin{aligned} &0 \leq p_{\mathbf{x}} \leq 1 \quad , \mathbf{x} = 1, 2, \ldots, |\mathbf{X}| \\ &\sum_{\mathbf{x}=1}^{|\mathbf{X}|} p_{\mathbf{x}} = 1 \end{aligned} \tag{2.2}$$

Thus the probability that a generic SBS stores a particular file c_j is given by

$$b_j = \sum_{\mathbf{x} \in \mathbf{X}_j} p_{\mathbf{x}} \tag{2.3}$$

where $\mathbf{X}_j \subseteq \mathbf{X}$ constitutes the set of all possible file combinations that contain the file c_j. The caching distribution $\mathbf{P} = \{p_{\mathbf{x}} : \mathbf{x} \in \mathbf{X}\}$ is considered as a key design parameter that controls the network performance as explained in the sequel. In this chapter, the uniform and Zipf caching are considered as the benchmarks file placement schemes. In the uniform caching, the M-files combinations are cached randomly and uniformly into the SBSs. Thus, $p_{\mathbf{x}} = \frac{1}{|\mathbf{X}|}$ and from (2.3), $b_j = \frac{\binom{J-1}{M-1}}{\binom{J}{M}} = \frac{M}{J}$. On the other hand, the Zipf caching follows the popularity of files, i.e., the Zipf distribution. Thus, the probability that a combination $\mathbf{x} \in \mathbf{X}$

to be stored at a generic SBS is $p_{\mathbf{x}} = \frac{1}{M} \sum_{j \in \mathbf{x}} a_j$.

2.2.3 Association Model

Without loss of generality, we assume that the indices of SBSs are ordered according to their distances from the typical user. The typical user is assumed to be served by the nearest SBS that stores the requested file. We also assume that the other users surrounding the typical user are traditionally associated to their geographically closest SBSs based on the maximum received power. It is assumed that each SBS picks a channel at random from the pool $\mathbf{S}$ to serve each of its associated users. A SBS does not assign the same channel to two of its associated users to avoid overwhelming intra-cell interference. Nonetheless, different SBSs can assign the same channel to one of their tagged users and thus inter-cell interference exists between BSs with scheduling ties. According to Slivnyak-Mecke theorem [62], evaluating the network performance for the typical user at the origin is sufficient and applicable to any generic location in the $2D$ plane. A realization of the considered system model is shown in Fig. 2.1. The red dot represents the typical user at the origin and the cyan dots are the other users. The red triangles represent the SBSs that store the requested file while the black squares are the others. The green triangles within the red dashed disk are the SBSs closer to the typical user than its serving SBS. The red and cyan dashed lines show the association policies of both the typical user and the other users in the network, respectively. The bold blue dashed lines represent the Voroni cells. The dashed

red circle encloses the SBSs that the file catering SBS should avoids their used channels when assigning channel to serve the typical user.

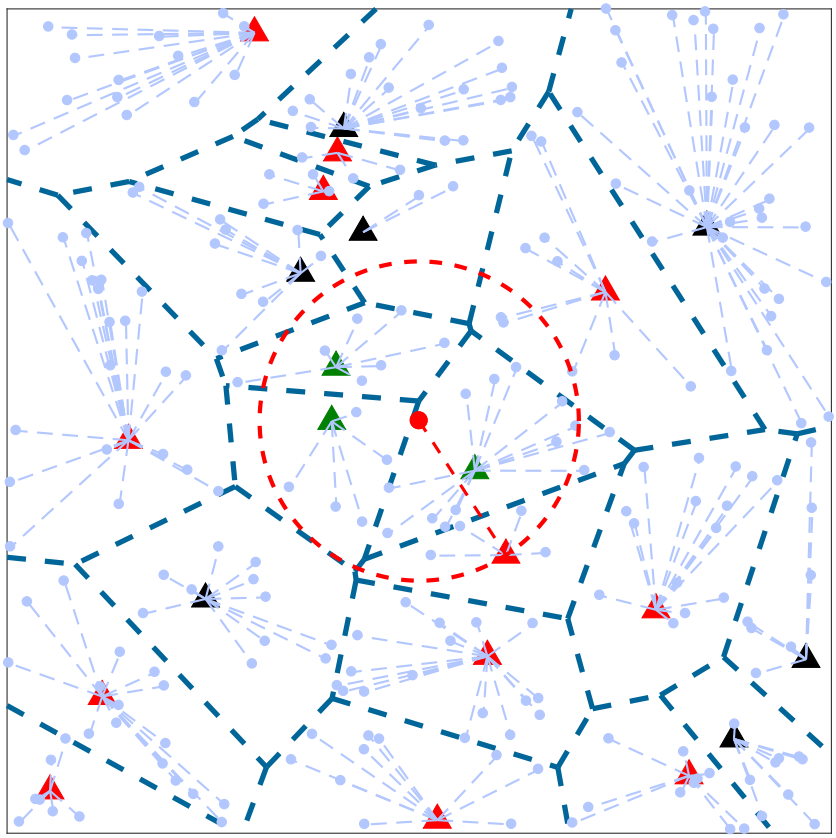


Figure 2.1: A realization of the network.

2.2.4 Performance Metric

The hit probability, defined by the joint event of having the requested file by the test user cached at an accessible edge SBS and its successful transmission from this SBS to the test user with a greater SINR than a certain threshold β, is the main performance metric considered in this chapter. According to the law of total

probability, the hit probability can be expressed as:

$$\mathcal{H}=\sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j(1-b_j)^{n-1}\left(\mathcal{C}_n\ \mathcal{O}_n+\mathcal{C}'_n\ (1-\mathcal{O}_n)\right), \tag{2.4}$$

where a_j is the probability that the test user requests the file c_j; $b_j\ (1-b_j)^{n-1}$ is the probability that the desired file is both available at the n^{th} SBS and is not stored in any of the $(n-1)$ closer SBSs; $\mathcal{C}_n$ denotes the coverage probability when OSA is successful, $\mathcal{C}'_n$ denotes the coverage probability when the OSA fails; and $\mathcal{O}_n$ represents the OSA success probability when the requesting user is served by the n^{th} SBS. According to (2.4), the hit event occurs when the requested file is stored in any of the SBSs covering the user. Otherwise, a miss event happens and the user is served by its geographically closest SBS, which downloads the requested file from the core network. This work explicitly focuses on the hit event analysis and optimization for both the unicast and multicast schemes. To confine the focus of the work to the hit probability and file placement optimization, the multicast overheads (e.g., synchronization overhead) are beyond the scope of this work.

2.3 OSA Success probability

The probability of free channel existence at the file catering n^{th} SBS is equivalent to the probability that the number of occupied channels by the $(n-1)$ SBSs that are closer to the typical user than its file catering SBS (i.e., within $\mathcal{D}(r_n)$) is less than the total number of available channels $|\mathbf{S}|$. The OSA success probability $\mathcal{O}_n$

is given by

$$\mathcal{O}_n = \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathcal{P}_{n-1}(\kappa) = 1 - \mathcal{P}_{n-1}(|\mathbf{S}|). \tag{2.5}$$

where $\mathcal{P}_{n-1}(\kappa)$ denotes the probability mass function (PMF) of the number of channels used by the SBSs within $\mathcal{D}(r_n)$. The first step to find $\mathcal{P}_{n-1}(\kappa)$ is to compute the PMF of the number of used channels by a generic SBS within $\mathcal{D}(r_n)$, which is equivalent to the number of users associated to it according to the adopted system model. Let U denote the number of users associated to a generic SBS, the PMF $\mathbb{P}[U = k]$ is given by [44, Appendix B]

$$\mathbb{P}[U = k] = \frac{\mathrm{v}^{\mathrm{v}}\,\Gamma(k+\mathrm{v})\,(q)^k}{\Gamma(\mathrm{v})\,\Gamma(k+1)\,(q+\mathrm{v})^{k+\mathrm{v}}} \tag{2.6}$$

where $q = \frac{\lambda_u}{\lambda_b}$ and $\mathrm{v} = 3.575$ is a constant related to the PPP Voronoi cell area distribution in $\mathbb{R}^2$. Note that this expression is valid only when the cellular users are assumed to be spatially distributed according to an independent PPP and their associations to the SBSs are based on the maximum average received signal power. Considering that each SBS randomly selects channels from the pool $\mathbf{S}$ to serve its tagged users, the event that two SBSs use the same channel might be occurred. Thus, the number of used channels by the $(n-1)$ SBSs within $\mathcal{D}(r_n)$ is not simply equivalent to the sum of their associated users. Consequently, the $\mathcal{P}_{n-1}(\kappa)$ is obtained recursively as in the following lemma.

Lemma 2.1 *The PMF of the number of used channels within $\mathcal{D}(r_n)$ can be given*

by the following recursive equation

$$\mathcal{P}_{n-1}(\kappa) = \sum_{v=0}^{\kappa} \mathcal{P}_{n-2}(v) \sum_{w=\kappa-v}^{\kappa} \mathbb{P}[U=w] \binom{w}{w-(\kappa-v)} \left(\frac{v}{|\mathbf{S}|}\right)^{w-(\kappa-v)} \left(1-\frac{v}{|\mathbf{S}|}\right)^{\kappa-v}, 0 \leq \kappa \leq |\mathbf{S}| \quad (2.7)$$

Proof. By using the induction method, conditioning that there are two SBSs within $\mathcal{D}(r_n)$ (i.e., $n=3$) and t channels used by one of them. Thus, the probability that κ channels are used within $\mathcal{D}(r_n)$ is given by

$$\mathcal{P}_2(\kappa \big| \, |\mathbf{S}_{bs1}| = t, t \leq \kappa) = \sum_{w=\kappa-t}^{\kappa} \mathbb{P}[U=w] \binom{w}{w-(\kappa-t)} \left(\frac{t}{|\mathbf{S}|}\right)^{w-(\kappa-t)} \left(1-\frac{t}{|\mathbf{S}|}\right)^{\kappa-t} \quad (2.8)$$

Therefore, by deconditioning on the number of used channels by the SBS $bs1$, we have

$$\mathcal{P}_2(\kappa) = \sum_{v=0}^{\kappa} \mathbb{P}[|\mathbf{S}_{bs_1}| = t] \sum_{w=\kappa-t}^{\kappa} \mathbb{P}[U=w] \binom{w}{w-(\kappa-t)} \left(\frac{t}{|\mathbf{S}|}\right)^{w-(\kappa-t)} \left(1-\frac{t}{|\mathbf{S}|}\right)^{\kappa-t} \quad (2.9)$$

Similarly, conditioning that there are three SBSs within $\mathcal{D}(r_n)$ and t channels used by SBSs bs_1 and bs_2. The probability that κ channels used within $\mathcal{D}(r_n)$ is

given by

$$\mathcal{P}_3(\kappa\big|\,|\mathbf{S}_{bs1}\cup\mathbf{S}_{bs2}|=t)=\sum_{w=\kappa-t}^{\kappa}\mathbb{P}[U=w]\binom{w}{w-(\kappa-t)}\left(\frac{t}{|\mathbf{S}|}\right)^{w-(\kappa-t)}\left(1-\frac{t}{|\mathbf{S}|}\right)^{\kappa-t} \tag{2.10}$$

Again by deconditioning on the number of used channels by SBSs bs_1 and bs_2, we have

$$\mathcal{P}_3(\kappa)=\sum_{v=0}^{\kappa}\mathcal{P}_2(t)\sum_{w=\kappa-t}^{\kappa}\mathbb{P}[U=w]\binom{w}{w-(\kappa-t)}\left(\frac{t}{|\mathbf{S}|}\right)^{w-(\kappa-t)}\left(1-\frac{t}{|\mathbf{S}|}\right)^{\kappa-t} \tag{2.11}$$

Therefore, by induction and deconditioning on the number of SBSs within $\mathcal{D}(r_n)$, the lemma is derived. ∎

2.4 The coverage Probability

The coverage probability is defined as the probability that the typical user can successfully achieve a specified SINR threshold β. The SINR distribution is characterized for the OSA success and failure cases, to find the coverage probabilities $\mathcal{C}_n$ and $\mathcal{C}'_n$, respectively.

2.4.1 Distance Distribution

The first step to analyze the SINR is to characterize the service distance distribution. Conditioning on the fact that the catering SBS is of order n, the probability

density function (PDF) of the service distance r_n is given by [64, Lemma 3]

$$f_{r_n}(r) = \frac{2\left(\pi\lambda_b r^2\right)^n}{r\Gamma(n)} e^{-\pi\lambda_b r^2} \ , 0 \leq r \leq \infty \tag{2.12}$$

According to the PPP, conditioning on r_n, the $(n-1)$ nearer SBS are uniformly and independently scattered in $\mathcal{D}(r_n)$. Hence, the distance distribution between the user and a randomly selected SBS of the $(n-1)$ nearer SBS is given by:

$$f_{r_i}(z|r_n) = \frac{2z}{r_n^2} \quad , 0 \leq z \leq r_n. \tag{2.13}$$

The distance distribution in (2.13) is needed to characterize interference when the OSA fails, i.e., when the test user experiences interference from $1 \leq \kappa \leq (n-1)$ SBSs within $\mathcal{D}(r_n)$.

2.4.2 SINR Distribution

Under the adopted system model, the SINR depends on whether a free channel exists at the catering file SBS or not. Firstly, we need to find the set of SBSs that use the same channel as the serving channel of the typical user. By exploiting the random channel selection policy at the SBSs and the independent thinning property of the PPP, the set of SBSs that interfere with the typical user forms a thinning PPP $\Psi_{b,\mathcal{T}}$ with intensity $\mathcal{T}\lambda_b$. The thinning parameter $\mathcal{T}$ is the probability that a generic SBS randomly selects a particular channel from $\mathbf{S}$, which is

given by [26]:

$$\mathcal{T} = 1 - \sum_{k=1}^{|\mathbf{S}|} \mathbb{P}[U = k] \frac{|\mathbf{S}| - k}{|\mathbf{S}|} \tag{2.14}$$

Then, the coverage probabilities $\mathcal{C}_n$ and $\mathcal{C}'_n$ are given by

$$\begin{aligned}
\mathcal{C}_n &= \mathbb{P}[SINR \geq \beta] = \mathbb{P}\big[\frac{P h_n r_n^{-\eta}}{\sigma_n^2 + \sum_{y_i \in \Psi_{b,\mathcal{T}}, r_i > r_n} P h_i r_i^{-\eta}} \geq \beta\big] \\
&= \mathbb{P}\big[\frac{P h_n r_n^{-\eta}}{\sigma_n^2 + \mathcal{I}_{out}} \geq \beta\big] \\
\mathcal{C}'_n &= \mathbb{P}\big[\frac{P h_n r_n^{-\eta}}{\sigma_n^2 + \sum_{y_i \in \Psi_{b,\mathcal{T}}, r_i < r_n} P h_i r_i^{-\eta} + \mathcal{I}_{out}} \geq \beta\big] \\
&= \mathbb{P}\big[\frac{P h_n r_n^{-\eta}}{\sigma_n^2 + \mathcal{I}_{in} + \mathcal{I}_{out}} \geq \beta\big]
\end{aligned} \tag{2.15}$$

where h_n(resp. h_i) is the channel gain between the typical user and its file catering SBS (resp. the i^{th} interfering SBS). r_n(resp. r_i) is the Euclidean distance between the typical user and its file catering SBS (resp. the i^{th} interfering SBS). σ_n^2 is the serving channel noise. $\mathcal{I}_{in}$ and $\mathcal{I}_{out}$ represent the aggregate interference of the SBSs closer to and farther to the typical user from its file catering SBS, respectively.

Conditioning on the distance r_n between the typical user at the origin and its file catering SBS, the conditional coverage probabilities are given by

$$\begin{aligned}
\mathcal{C}_n(r_n) &= e^{\frac{-\beta r_n^{\eta} \sigma_n^2}{P}} \ \mathcal{L}_{\mathcal{I}_{out}}\big(\frac{\beta r_n^{\eta}}{P}\big) \\
\mathcal{C}'_n(r_n) &= e^{\frac{-\beta r_n^{\eta} \sigma_n^2}{P}} \ \mathcal{L}_{\mathcal{I}_{out}}\big(\frac{\beta r_n^{\eta}}{P}\big) \ \mathcal{L}_{\mathcal{I}_{in}}\big(\frac{\beta r_n^{\eta}}{P}\big)
\end{aligned} \tag{2.16}$$

where $\mathcal{L}_{\mathcal{I}}(t) = \mathbb{E}[e^{-t\mathcal{I}}]$ denotes the Laplace transform (LT) of $\mathcal{I}$. $\mathcal{L}_{\mathcal{I}_{in}}$ and $\mathcal{L}_{\mathcal{I}_{out}}$ are the Laplace transforms (LTs) of $\mathcal{I}_{in}$ and $\mathcal{I}_{out}$, respectively.

the LT of the interference from the SBSs outside $\mathcal{D}(r_n)$ can be obtained similar

to [65] with considering only the SBSs that use the same channel of the typical user. Thus, $\mathcal{L}_{\mathcal{I}_{out}}$ is given by

$$\mathcal{L}_{\mathcal{I}_{out}}\left(\frac{\beta r_n^{\eta}}{P}\right) = \exp\left(- \pi \mathcal{P} \lambda_b r_n^2 \beta^{\frac{2}{\eta}} \int_{z=\beta^{\frac{-2}{\eta}}}^{\infty} \frac{1}{1+z^{\frac{\eta}{2}}} dz \right) \tag{2.17}$$

On the other hand, the LT $\mathcal{L}_{\mathcal{I}_{in}}$ of the interference from the SBSs within $\mathcal{D}(r_n)$ at the OSA failure is given by the following Lemma

Lemma 2.2 *The LT of the interference from the SBSs within $\mathcal{D}(r_n)$ is given by*

$$\mathcal{L}_{\mathcal{I}_{in}}\left(\frac{\beta r_n^{\eta}}{P}\right) = \left(\frac{2\mathcal{T}}{r_n^2} \int_{r=0}^{r_n} \frac{1}{1+\beta r_n^{\eta}\ r^{-\eta}} r dr + 1 - \mathcal{T}\right)^{n-1} - (1-\mathcal{T})^{n-1} \tag{2.18}$$

Proof. In the scenario of no available free channel, the typical user is served over a randomly chosen channel by the file catering SBS. Thus, the number of inside r_n interfering SBSs κ is a random variable with a maximum value $(n-1)$. The PMF of κ is given by

$$\mathbb{P}[\kappa = a] = \binom{n-1}{a} \mathcal{T}^a\ (1-\mathcal{T})^{n-a-1} \quad , 1 \leq a \leq (n-1) \tag{2.19}$$

Given that the number of the inside r_n interfering SBSs $\kappa = a$, the LT $\mathcal{I}_{in,\kappa}$ is

given by

$$\begin{aligned}\mathcal{L}_{\mathcal{I}_{in,\kappa}}(t) &= \mathbb{E}_{\mathcal{I}_{in},\kappa}\left[e^{-t\mathcal{I}_{in}}|\kappa=a\right] \\ &= \mathbb{E}_{\kappa}\left[\mathbb{E}_{y_i\in\Psi_{b,\kappa},r_i<r_n}\left[e^{-t\sum_i Ph_i r_i^{-\eta}}\right]\ |\kappa=a\right] \\ &\overset{(a)}{=} \mathbb{E}_{\kappa}\left[\mathbb{E}_{r_{n-1}}\left[\mathbb{E}_h[e^{-tPhr^{-\eta}}\ |\ r=r_n]\right]^{\kappa}|\kappa=a\right] \\ &\overset{(b)}{=} \mathbb{E}_{\kappa}\left[\left[\int_{r=0}^{r_n}\frac{1}{1+tPr^{-\eta}}\ f_{r_{n-1}}(r|r_n)dr\right]^{\kappa}|\kappa=a\right] \end{aligned} \tag{2.20}$$

The equality (a) is by the fact that the PPP consists of uniformly distributed nodes, so intuitively the conditional PDF of the κ nodes on the n^{th} node is equivalent to the κ times the conditional PDF of two consecutive nodes. Equality (b) from the channel gain exponential distribution with averaging over the conditional PDF $f_{r_{n-1}}(r|r_n)$ that is given in (2.13). Replacing t in the above equation with $(\frac{\beta r_n^{\eta}}{P})$ and averaging over the random variable κ, the unconditional LT $\mathcal{L}_{\mathcal{I}_{in}}$ can be obtained. Finally, by applying the binomial theorem, the lemma can easily verified. ∎

By averaging over the PDF of the serving distance r_n between the typical user at the origin and its file catering SBS, $f_{r_n}(r)$ in (2.12), the unconditional coverage probabilities are given by (2.21).

In the special case of interference-limited network (i.e., $\sigma_n^2 = 0$) and path loss exponent $\eta = 4$ (which is common for wireless networks), and with integral manipulations, the unconditional coverage probabilities in (2.21) turn to the simple closed-form expressions explained in the following corollary

$$
\begin{aligned}
\mathcal{C}_n &= \frac{2(\pi\lambda_b)^n}{\Gamma(n)} \int_0^{\infty} v^{2n-1} e^{\frac{-\beta v^{\eta}\sigma_n^2}{P}} \exp\left(-\pi\mathcal{T}\lambda_b v^2 \beta^{\frac{2}{\eta}} \int_{z=\beta^{\frac{-2}{\eta}}}^{\infty} \frac{1}{1+z^{\frac{\eta}{2}}} dz \right) e^{-\pi\lambda_b v^2} dv \\
\mathcal{C}'_n &= \frac{2(\pi\lambda_b)^n}{\Gamma(n)} \int_0^{\infty} v^{2n-1} e^{\frac{-\beta v^{\eta}\sigma_n^2}{P}} \exp\left(-\pi\mathcal{T}\lambda_b v^2 \beta^{\frac{2}{\eta}} \int_{z=\beta^{\frac{-2}{\eta}}}^{\infty} \frac{1}{1+z^{\frac{\eta}{2}}} dz \right) e^{-\pi\lambda_b v^2} \\
&\qquad \left\{ \left(1 - \mathcal{T} + \frac{2\mathcal{T}}{v^2} \int_{w=0}^{v} \frac{1}{1+\beta v^{\eta}\ w^{-\eta}} w dw \right)^{n-1} - (1-\mathcal{T})^{n-1} \right\} dv
\end{aligned}
\tag{2.21}
$$

---

corollary 2.1 *The closed-form expressions of the coverage probabilities for an interference-limited network and with $\eta = 4$ are given by*

$$
\begin{aligned}
\mathcal{C}_n &= \left(1 + \mathcal{T}\sqrt{\beta}\ \arctan(\sqrt{\beta}) \right)^{-n} \\
\mathcal{C}'_n &= \mathcal{C}_n \left[\left(1 - \mathcal{T}\sqrt{\beta}\ \arctan(\frac{1}{\sqrt{\beta}}) \right)^{n-1} - (1-\mathcal{T})^{n-1} \right]
\end{aligned}
\tag{2.22}
$$

Finally, by substituting from (2.21) and (2.22) into (2.4), we end up with the general and closed-form expressions of the hit probability for the proposed multi-channel system, respectively.

It is worth noting that the hit probability of the single-channel cache-enabled network, which is widely used in the literature can be derived from the multi-channel system. In the single-channel scenario, all the SBSs use the same channel to serve their associated users. Therefore, the typical user suffers from interference from all SBSs except its file catering SBS. Following the previous analysis, $\mathcal{T}$ is omitted and k has a deterministic with value $(n-1)$. Thus, the single-channel system hit probability $\mathcal{H}_{\text{single}}$ is given by the following Lemma.

Lemma 2.3 *The hit probability of the single-channel system is given by*

$$\mathcal{H}_{single} = \sum_{j\in\mathcal{J}} a_j \sum_{n=1}^{\infty} b_j(1-b_j)^{n-1}\ \mathcal{C}_{single}$$

$$\mathcal{C}_{single} = \frac{2(\pi\lambda_b)^n}{\Gamma(n)} \int_0^{\infty} v^{2n-1} \left(\frac{2}{v^2} \int_{w=0}^{v} \frac{1}{1+\beta v^{\eta}\ w^{-\eta}} w dw \right)^{n-1} \exp\left(-\pi\lambda_b v^2 \beta^{\frac{2}{\eta}} \int_{z=\beta^{\frac{-2}{\eta}}}^{\infty} \frac{1}{1+z^{\frac{\eta}{2}}} dz \right) e^{\frac{-\beta v^{\eta}\sigma_n^2}{P}} e^{-\pi\lambda_b v^2} dv \qquad (2.23)$$

Again, in the special case of interference-limited network and $\eta = 4$, the coverage probability in (2.23), we end up with a closed-form expression that is given by

$$\mathcal{C}_{\text{single}} = \left(1+\sqrt{\beta}\ \arctan(\sqrt{\beta})\right)^{-n} \left(1-\sqrt{\beta}\ \arctan(\frac{1}{\sqrt{\beta}})\right)^{n-1} \qquad (2.24)$$

2.5 Numerical results

In this section, the simulation for the adopted system is held to verify the derived mathematical paradigm. The parameters $\lambda_b = 4$ SBSs/Km^2 and $\lambda_u = 40$ SBSs/Km^2 were chosen to conduct the analysis. All the SBSs have the same transmitted power $P = 1$ watt. The path loss exponent $\eta = 4$ is considered. The assumed Zipf caching is compared with the uniform caching in which the popular files are distributed into the SBSs randomly irrespective to their popularity. Also, the proposed system is compared with the single channel system to reveal the improvement in the achieved hitting probability.

Fig. 2.2 illustrates the hit probability versus the SINR threshold for both

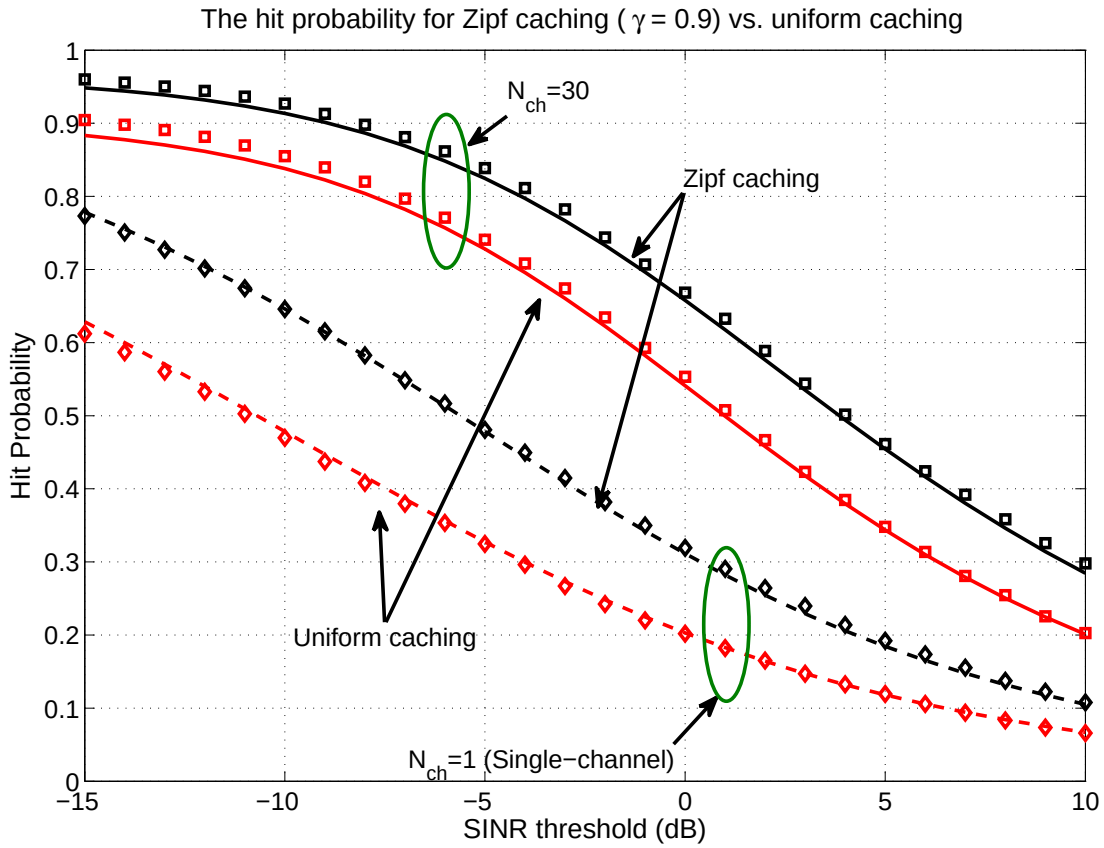


Figure 2.2: The hit probability vs. the SINR with $J = 15$ & $K = 5$ & $\frac{\lambda_u}{\lambda_b} = 10$ and $c_j = 1$

the multichannel and the single-channel systems using uniform and Zipf caching schemes considering the requested file is $c_j = 1$. A clear match can be noticed between our derived expressions and the simulation results, for both single and multi-channel systems. Furthermore, Fig.2.2 quantifies the gain of the multi-channel system over the single-channel case and reveals the significant improvement of system performance using the Zipf caching compared with the uniform caching scheme. However, focusing on the first popular file of the highest popularity a_1 results in an unfair study.

Therefore, Fig.2.3 shows the hit probability versus the requested file index. It

can be observed that the hitting probability is independent from the file index for uniform caching scheme. This is because the uniform caching scheme distributes the files into SBSs randomly and uniformly irrespective to their popularity. On the other hand, the hit probability of the Zipf caching scheme highly depends on the requested file popularity. The hitting probability dramatically degrades when the typical user requests any other files rather than the first one in the case of high Zipf parameter. Moreover, Fig.2.3 indicates that the multi-channel system always outperforms the single-channel scenario. In addition, the uniform caching slightly outperforms the Zipf caching when the files with low popularity are requested.

Fig.2.4 illustrates the hit probability versus the cache size. It can be noticed

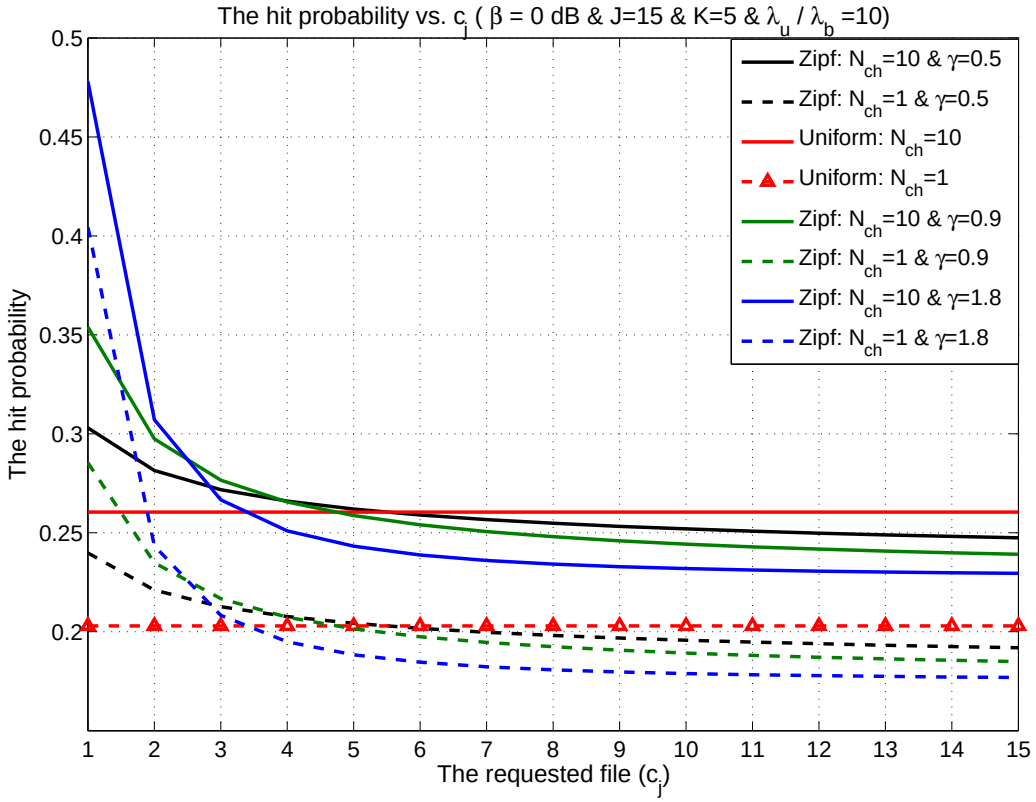


Figure 2.3: The hit probability vs. the requested file c_j with $J = 15$ & $K = 5$ & $\frac{\lambda_u}{\lambda_b} = 10$ at different values of Zipf parameter.

that the performance of the system improves as the cache size increases. This highlights the tradeoff between the cache-enabled network performance and the network storage resources. It can be observed that the hit probability behaves even linearly or nonlinearly with the cache size. It can be explained by the fact that at most m SBSs can have signal-to-interference ratio (SIR) greater than $\frac{1}{m}$ for any positive integer m [66]. Thus, there are at most 10 SBSs that have SIR of -10 dB and only one SBS with SIR of 0 dB. This fact explains the non-linearity of the hit probability at -10 dB scenario for all the displayed number of channels. On the other hand, almost all the trends are linear at the 0 dB scenario. The dependency on the number of channels comes from the fact that increasing the

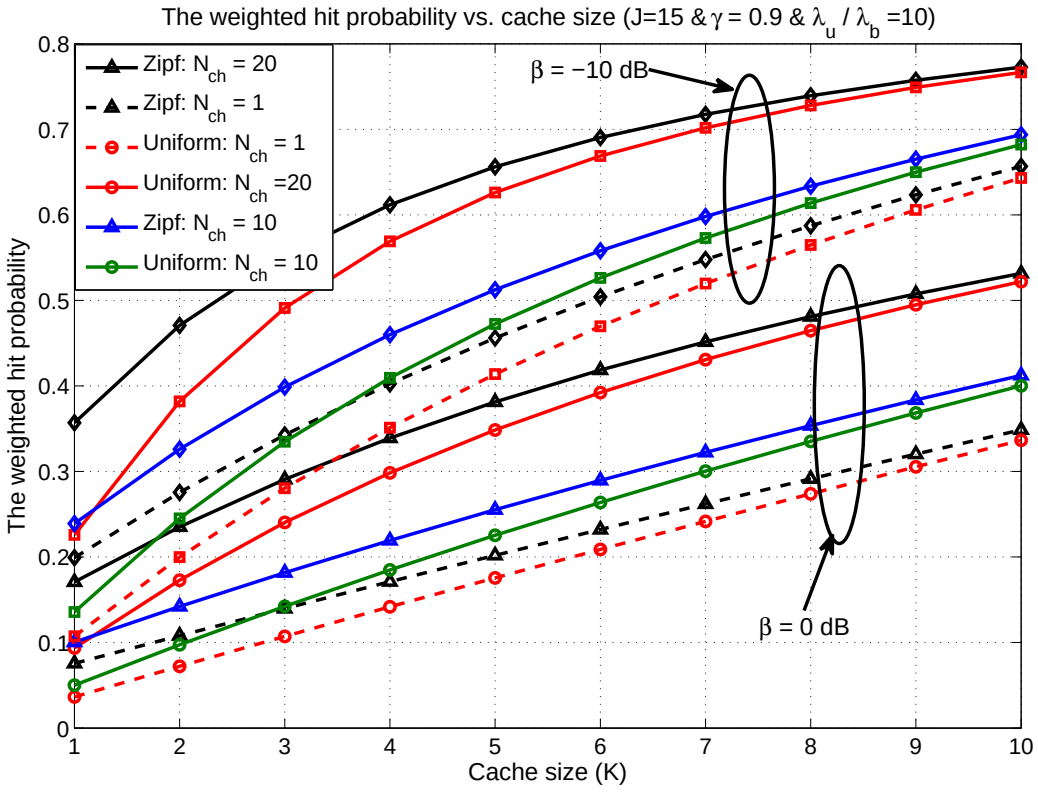


Figure 2.4: The hit probability vs. the cache size K with $\beta = 0$ dB& $J = 15$ & $\gamma = 0.9$ & $\frac{\lambda_u}{\lambda_b} = 10$ at different number of channels.

number of channels enhances the chance that typical user is served by multiple SBSs and this combination ends up with the non-linear trend.

Fig.2.5 shows the hit probability versus the Zipf parameter γ. The figure indicates that the hit probability increases as the Zipf parameter increases. It can be noticed that the multi-channel system has better performance than the single-channel system. In addition, the Zipf caching outperforms the uniform caching scenario thanks to the opportunistic spectrum access and the caching diversity.

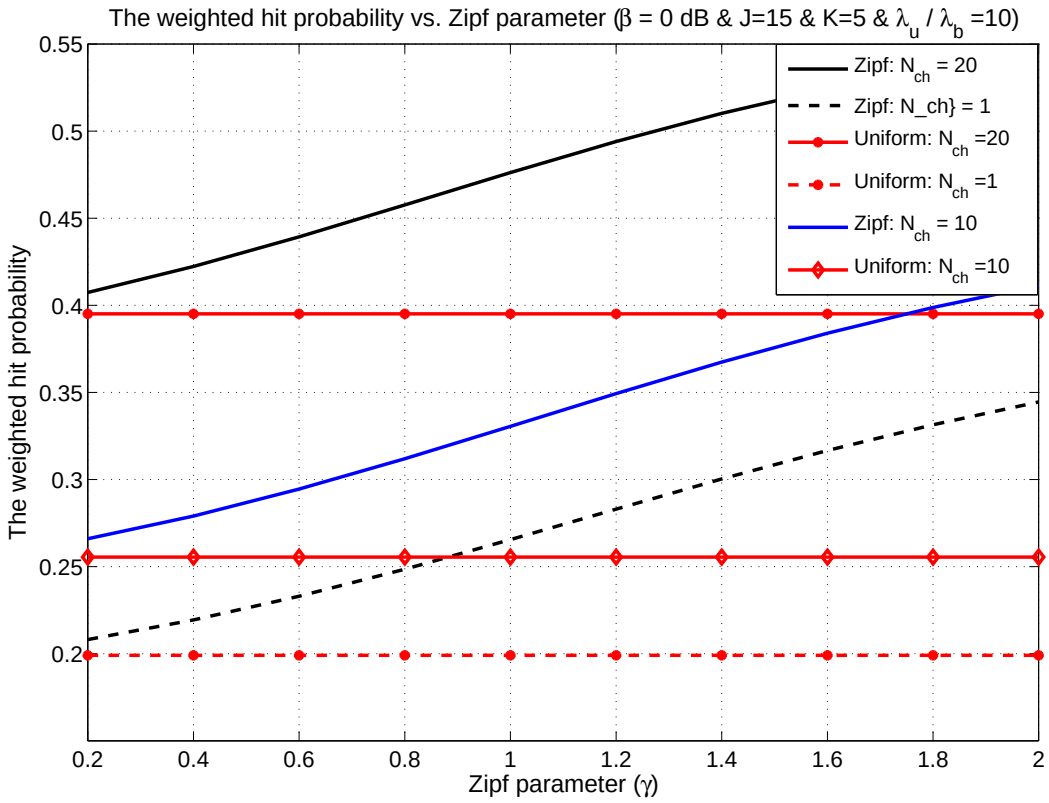


Figure 2.5: The hit probability vs. the Zipf parameter γ with $\beta = 0$ dB& $J = 15$ & $K = 5$ & $\frac{\lambda_u}{\lambda_b} = 10$ at different number of channels.

2.6 Chapter Summary

This chapter considers an OSA approach in multi-channel environment for single-user cache-enabled network. A tractable mathematical model for the hit probability is developed, which is reduced to closed form expression in special cases. The numerical results both verify the derived analytical paradigms and quantify the gains of the considered multi-channel model. We end up by the achieved improvement of the joint multi-channel cached-enabled network over the single-channel system.

CHAPTER 3

OPTIMAL CACHING IN 5G NETWORKS WITH OPPORTUNISTIC SPECTRUM ACCESS

3.1 Introduction

In this chapter, we extend the work in chapter 2 to the full cache-enabled system in which all the users in the network request popular contents. We develop a mathematical model, based on the stochastic geometry [25], for cache-enabled 5G networks with OSA in a multi-channel environment. We assume a PPP cellular network with Zipf distributed file popularity. The PPP assumption is widely accepted for modeling cellular networks to retain tractability [26, Table I] and has

been verified in [27–29] by several empirical studies. In contrast to the literature, we consider both caching design and OSA based interference management for unicast and multicast content delivery. Since the requested file by any user may not be served from the geographically nearest SBS, the catering SBS utilizes OSA to alleviate interfering with all SBSs that are closer to this demanding user. The SBSs that are closer to the user than the serving SBS can be determined through the *neighboring cell list* (NCL) reported by each user. The NCL contains the IDs of candidate BSs for association, and is mainly used for handover management and user-centric coordinated/cooperative multi-BS transmissions [67–70]. The serving SBS can inquire about the channels used by each of the SBSs closer to the demanding user on the signaling backhaul link. To this end, we characterize the hit probability and propose a simple concave approximation for it. This concave approximate expression is then used to formulate and solve a reduced-complexity file placement optimization problem, and to derive a closed-form solution for it. This proposed solution is benchmarked against the conventional Zipf and uniform based caching.

The rest of this chapter is organized as follows: Section 3.2 presents the proposed system model. Section 3.3 derives expressions for the OSA success probability for both the unicast and multicast transmission schemes. The coverage probabilities under the OSA success and fail are obtained in Section 3.4. Section 3.5 formulates the optimization problem to find the sub-optimal caching distribution that maximizes the hit probability for both the unicast and multicast

schemes. Section 3.6 introduces the numerical results that verify the developed mathematical framework in the chapter.

3.2 System Model

This section explains the proposed system model for the full cache-enabled network. We adopt the network and caching models that are used in chapter 2. However, the channel assignment policy is changed. In addition, two transmission schemes (i.e., the unicast and multicast) are adopted as illustrated below.

3.2.1 Channel Assignment Policy

Each user requesting file c_j is associated to the nearest SBS that caches a files combination $\mathbf{x} \in \mathbf{X}_j$ that includes the requested file c_j. Since $M < J$, the catering SBS is not necessarily the geographically closest SBS. When the user is served by the n^{th} SBS, the requested file is transmitted over a channel that is not used by the $(n-1)$ closer SBSs. Particularly, the catering SBS avoids interfering with any of the $(n-1)$ SBSs that exist within $\mathcal{D}(r_n)$. If the SBSs within $\mathcal{D}(r_n)$ utilize the complete set of channels for serving their attached users, the catering SBS randomly and uniformly selects a channel to transmit the requested file to the test user. Again we refer to the former (latter) event, representing the presence (absence) of an unused channel for the catering SBS transmission, as OSA success (failure).

3.2.2 Transmission Scheme

In this work, we develop a framework for networks adopting both the *unicast-based* and *multicast-based* transmission modes. In the former scenario, the SBSs serve each of its associated users over a different channel from the set $\mathbf{S}$. On the other side, multicast-based transmission mode groups all users that request the same file and serve them all over the same channel. As highlighted in Section 1, each transmission scheme has its own merits, and hence, both are utilized in practice. Since multicast transmission aggregates similar requests on the same channel, at most M channels are used by an SBS (i.e., when all of its cached files are requested). The effect of user blocking when the number of users requesting files from a SBS is greater than the number of available channels $|\mathbf{S}|$ is beyond the scope of this work.

Again we consider the hit probability as the main performance metric of the system. For reading convenience, we recall its expression that is introduced in chapter 2, which is expressed as

$$\mathcal{H}=\sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j(1-b_j)^{n-1} \left(\mathcal{C}_n \, \mathcal{O}_n+\mathcal{C}'_n \, (1-\mathcal{O}_n)\right), \tag{3.1}$$

In the next subsection, we derive expressions for the OSA success and the coverage probabilities under both the unicast and multicast transmission schemes.

3.3 Opportunistic Spectrum Access

This section derives the OSA probabilities for unicast and multicast transmission schemes. The OSA success probability $\mathcal{O}_n$ is defined in (4.7) and is used to compute the weighted average of the coverage probabilities in the considered caching systems. As stated earlier, the OSA process is considered successful if the catering SBS of order n finds a free channel, i.e., a channel not used by any of the $(n-1)$ SBSs closer to the requesting user. Hence, the OSA success depends on three factors: i) the number of users requesting each file, ii) the order of the catering SBS (i.e., n), and iii) the channel assignment scheme per SBS for the users requesting each file. The first two factors are independent from the transmission schemes. In contrast, the channel assignment is affected by whether a unicast or multicast transmission mode is adopted. Note that the correlations among the adjacent Voronoi cell sizes along with the OSA impose an interdependence between the number of used channels in adjacent SBSs, which impedes the exact characterization for OSA. Hence, we resort to the following two approximations:

Approximation 3.1 *Independent and identical uniform scheduling for users at each SBS is assumed, which ignores the OSA scheduling interactions among adjacent SBSs.*

Approximation 3.2 *The numbers of users requesting each file from a common SBS are assumed to be independent. Furthermore, the numbers of users served by adjacent SBSs are assumed to be independent.*

Note that Approximations 3.1 and 3.2 are mandatory for tractability and are validated in Section-3.6 through the matching between the simulation and the analytical results.

Let U_i denote the number of users that request c_i and are associated to the same SBS that stores it. Using the thinning property of the PPP [71], the complete set of users requesting the file c_i constitutes a thinned PPP $\mathbf{\Phi}_{u,i}$ of intensity $a_i\lambda_u$ and the set of SBSs that store c_i forms an independent thinned PPP $\mathbf{\Phi}_{b,i}$ of intensity $b_i\lambda_b$. Exploiting Approximation 3.2, the PMF of U_i is given by [44, Appendix B]

$$\mathbb{P}[U_i = k] = \frac{\mathrm{v}^{\mathrm{v}}\,\Gamma(k+\mathrm{v})\,(q_i)^k}{\Gamma(\mathrm{v})\,\Gamma(k+1)\,(q_i+\mathrm{v})^{k+\mathrm{v}}} \tag{3.2}$$

where $q_i = \frac{a_i\lambda_u}{b_i\lambda_b}$ and $\mathrm{v} = 3.575$ is a constant related to the PPP Voronoi cell area distribution in $\mathbb{R}^2$. Note that this expression is valid only when the cellular users are assumed to be spatially distributed according to an independent PPP and their associations to the SBSs are based on the maximum average received signal power, i.e., each user is associated with the nearest SBS that stores its requested file.

The number of users requesting each file is independent from the transmission scheme. Hence, (3.2) is legitimate to compute the OSA success probability in both the unicast and multicast schemes, which are presented in the sequel.

3.3.1 Unicast Mode

Exploiting the fact that each user is served on a unique channel, the PMF of the number of channels used by a generic SBS is given by the following lemma.

Lemma 3.1 *Let r_n be the distance between the catering SBS and the requesting user, then the PMF of the number of used channels by a generic SBS inside r_n, denoted as N_{-j}, is given by:*

$$\mathbb{P}[N_{-j}=k]=\sum_{\mathbf{x}\in\mathbf{X}_{-j}}\frac{p_{\mathbf{x}}}{1-b_j}\sum_{\kappa_a=0}^{\min(k,M)}\sum_{\{\mathbf{A}\subseteq\mathbf{x}:\ |\mathbf{A}|=\kappa_a\}}\left\{\mathbb{P}\left[\sum_{i\in\mathbf{A}}U_i=k\right]\prod_{m\in\mathbf{x}\setminus\mathbf{A}}\mathbb{P}[U_m=0]\right\}$$
$$,0\leq k\leq|\mathbf{S}| \qquad (3.3)$$

where the subscript $(-j)$ indicates that the requested file c_j is not stored by any of the SBSs within $\mathcal{D}(r_n)$.

Proof. The number of channels used by a generic SBS inside r_n is the number of users associated to it and request any of its stored files. Note that the cached combination at any SBS closer to the test user than its serving SBS (inside r_n) does not contain the requested file c_j, i.e., $x\in\mathbf{X}_{-j}$. Considering that might be some files that are not requested, $0\leq k_a\leq M$ denotes the number of active files (i.e., the files that are requested). Thus, the conditional PMF of the number of used channels by this SBS is given by

$$\mathbb{P}[N_{-j}=k\Big|\text{SBS stores a combination }\mathbf{x}\in\mathbf{X}_{-j}]$$
$$=\sum_{\kappa_a=0}^{\min(k,M)}\sum_{\{\mathbf{A}\subseteq\mathbf{x}:\ |\mathbf{A}|=\kappa_a\}}\Big\{\mathbb{P}[\sum_{i\in\mathbf{A}}U_i=k]\prod_{m\in\mathbf{x}\setminus\mathbf{A}}\mathbb{P}[U_m=0]\Big\}\quad ,0\leq k\leq|\mathbf{S}| \quad (3.4)$$

The probability that a SBS stores a combination $\mathbf{x} \in \mathbf{X}_{-j}$ is $\frac{p_{\mathbf{x}}}{(1-b_j)}$. Therefore, by averaging over all possible combinations in the set $\mathbf{X}_{-j}$, the unconditional PMF of N_{-j} is obtained. ∎

Recalling that U_i is the number of users that request c_i, the term $\mathbb{P}\left[\sum_{i\in\mathbf{A}} U_i = k\right]$ in (3.3) denotes the PMF of the total number of users requesting the files in the set $\mathbf{A}$. Due to unicast, the number of users is equal to number of used channels. According to (3.2), the PMF of each of these random variables depends on the file popularity a_i and caching probability b_i. Exploiting Approximation 3.2, the PMF of $\sum_{i\in\mathbf{A}} U_i$ in (3.3) is approximated via the convolution of the probability mass functions (PMFs) of $U_i \; \forall \; i \in \mathbf{A}$. Alternatively, the PMF of the sum can be defined by the product of the characteristic functions. Let ϕ_{u_i} and ϕ_u denote the characteristic function of U_i and $\sum_{i\in\mathbf{A}} U_i$, respectively. Thus, ϕ_{u_i} is given by

$$\phi_{u_i}(t) = \mathbb{E}[e^{\mathrm{i}t\kappa}] = \sum_{\kappa=0}^{\infty} e^{\mathrm{i}t\kappa}\, \mathbb{P}[U_i = \kappa] = \left(\frac{\mathrm{v}\, q_i^{-1}}{1 - e^{\mathrm{i}t} + \mathrm{v}\, q_i^{-1}}\right)^{\mathrm{v}} \tag{3.5}$$

where $\mathrm{i} = \sqrt{-1}$. Therefore, the characteristic function ϕ_u is given by

$$\phi_u(t) = \prod_{i\in\mathbf{A}} \left(\frac{\mathrm{v}\, q_i^{-1}}{1 - e^{\mathrm{i}t} + \mathrm{v}\, q_i^{-1}}\right)^{\mathrm{v}} \tag{3.6}$$

Hence, the PMF $\mathbb{P}\left[\sum_{i\in\mathbf{A}} U_i = k\right]$ can be obtained from ϕ_u as follows [72, Section 4.5.2]

$$\mathbb{P}\left[\sum_{i\in\mathbf{A}} U_i = k\right] = \lim_{z\to\infty} \frac{1}{2z} \int_{-z}^{z} e^{-\mathrm{i}tk} \phi_u(t) dt \tag{3.7}$$

While (3.3) describes the number of channels used by a single SBS that does not contain c_j, the OSA process requires calculating the total number of unique channels (i.e., without counting repetitions of the same channel) used by all SBSs within $\mathcal{D}(r_n)$. Let $\mathcal{P}_{n-1}(\kappa)$ denote the PMF of the number of channels used by the SBSs within $\mathcal{D}(r_n)$. Exploiting the i.i.d. uniform scheduling approximation (i.e., Assumption 3.1) and following [73], $\mathcal{P}_{n-1}(\kappa)$ can be expressed via the following recursive equation:

$$\mathcal{P}_{n-1}(\kappa) \approx \sum_{v=0}^{\kappa} \mathcal{P}_{n-2}(v) \sum_{w=\kappa-v}^{\kappa} \Bigg\{ \mathbb{P}[N_{-j}=w] \binom{w}{w-(\kappa-v)} \left(\frac{v}{|\mathbf{S}|}\right)^{w-(\kappa-v)} \left(1-\frac{v}{|\mathbf{S}|}\right)^{\kappa-v} \Bigg\}; \quad 0 \leq \kappa \leq |\mathbf{S}| \quad (3.8)$$

Using (3.3)-(4.11), the OSA success probability can be expressed as:

$$\mathcal{O}_n = \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathcal{P}_{n-1}(\kappa) = 1 - \mathcal{P}_{n-1}(|\mathbf{S}|). \quad (3.9)$$

3.3.2 Multicast Mode

In the multicast mode, the users are grouped based on their demands, such that matching demands are served on the same channel. Hence, the number of channels per SBS, which is upper bounded by M, is a function of the popularity of the cached files. Accordingly, the PMF of the number of used channels by a generic

SBS inside $\mathcal{D}(r_n)$ is given by:

$$\mathbb{P}[N_{-j}=\kappa]=\sum_{\mathbf{x}\in\mathbf{X}_{-j}}\frac{p_{\mathbf{x}}}{1-b_j}\sum_{\{\mathbf{A}\subseteq\mathbf{x}:\ |\mathbf{A}|=\kappa\}}\left[\prod_{m\in\mathbf{A}}\zeta_m\prod_{m\in\mathbf{x}\setminus\mathbf{A}}(1-\zeta_m)\right],\quad 0\leq\kappa\leq M, \tag{3.10}$$

where ζ_j denotes the probability that file c_j is requested from a generic SBS. Particularly, ζ_j is equal to the probability that at least one of the SBS's associated users requests c_j. Using (3.2), ζ_j is given by:

$$\zeta_j=\mathbb{P}[U_j\geq 1]=1-\left[\mathrm{v}^{-1}\frac{a_j\,\lambda_u}{b_j\,\lambda_b}+1\right]^{-\mathrm{v}}, \tag{3.11}$$

Exploiting Assumption 3.1, and restricting the number of channels used by each SBS to the cache size (i.e., M), the recursive equation for the number of channels used by SBSs inside $\mathcal{D}(r_n)$ in (4.11) becomes:

$$\mathcal{P}_{n-1}(\kappa)\approx\sum_{v=0}^{\min\left(\kappa,(n-2)M\right)}\mathcal{P}_{n-2}(v)\sum_{w=\kappa-v}^{\kappa}\left[\mathbb{P}[N_{-j}=w]\binom{w}{w-(\kappa-v)}\left(\frac{v}{|\mathbf{S}|}\right)^{w-(\kappa-v)}\left(1-\frac{v}{|\mathbf{S}|}\right)^{\kappa-v}\mathbb{1}_{\{w\leq M\}}\right];\quad 0\leq\kappa\leq\min\Big(|\mathbf{S}|\,,M(n-1)\Big) \tag{3.12}$$

where $\mathbb{1}_{\{.\}}$ is the indicator function, i.e., $\mathbb{1}_{\{z\}}=1$ if condition z is true, and is zero otherwise.

3.4 Coverage Probability

In this section, we find the coverage probabilities for both the unicast and multicast transmission modes. The coverage probability is defined as the probability that the test user can successfully achieve a specified SINR threshold β. The SINR distribution is characterized for the OSA success and failure cases, to find the coverage probabilities $\mathcal{C}_n$ and $\mathcal{C}'_n$, respectively.

3.4.1 SINR analysis

The SINR distribution depends on whether the OSA succeeded or not. Let $\tilde{\mathbf{\Phi}}_b \subset \mathbf{\Phi}_b$ denote the set of interfering SBSs that are using the same channel allocated to the test user. OSA success implies that the disc $\mathcal{D}(r_n)$ is interference free (i.e., $\tilde{\mathbf{\Phi}}_b \cap \mathcal{D}(r_n) = \emptyset$). Hence, the coverage probability is expressed as:

$$\mathcal{C}_n = \mathbb{P}\left[\frac{Ph_n r_n^{-\eta}}{\sigma_n^2 + \mathcal{I}_{out}} \geq \beta\right] = \mathbb{P}\left[\frac{Ph_n r_n^{-\eta}}{\sigma_n^2 + \sum\limits_{y_i \in \tilde{\mathbf{\Phi}}_b} Ph_i \left\|y_i\right\|^{-\eta}} \geq \beta\right] \tag{3.13}$$

where h_n(resp. h_i) is the channel gain between the test user and its catering SBS (resp. the i^{th} interfering SBS), $\|\cdot\|$ is the Euclidean norm, $\mathcal{I}_{out} = \sum_{y_i \in \tilde{\mathbf{\Phi}}_b} Ph_i \left\|y_i\right\|^{-\eta}$ is the total interference power resulting from the SBSs outside $\mathcal{D}(r_n)$ using the same channel allocated to the test user, and σ_n^2 is the noise variance. Let $\mathcal{T}$ be the probability that a generic SBS outside $\mathcal{D}(r_n)$ is using the same channel as the test user. Exploiting Assumption 3.1 along with the thinning property of the PPP, the set of interfering SBSs $\tilde{\mathbf{\Phi}}_b$ forms an independent PPP

with intensity $\mathcal{T}\lambda_b$.

On the other hand, OSA failure implies that there exist interfering SBSs within $\mathcal{D}(r_n)$, which should be considered in the coverage probability analysis. Consequently, the coverage probability becomes:

$$\mathcal{C}'_n = \mathbb{P}\left[\frac{Ph_n r_n^{-\eta}}{\sigma_n^2 + \mathcal{I}_{in} + \mathcal{I}_{out}} \geq \beta\right], \tag{3.14}$$

where $\mathcal{I}_{out} = \sum_{y_i \in \tilde{\mathbf{\Phi}}_b \setminus \mathcal{D}(r_n)} Ph_i \left\|y_i\right\|^{-\eta}$ is the aggregated interference power resulting from the SBSs outside $\mathcal{D}(r_n)$ using the serving channel of the test user, and $\mathcal{I}_{in} = \sum_{y_i \in \tilde{\mathbf{\Phi}}_b \cap \mathcal{D}(r_n)} Ph_i \left\|y_i\right\|^{-\eta}$ is the total interference power resulting from the SBSs inside $\mathcal{D}(r_n)$ using the same channel allocated to the test user. Let $\mathcal{T}_{-j}$ be the probability that an SBS in $\tilde{\mathbf{\Phi}}_b \cap \mathcal{D}(r_n)$ is using the same channel as the test user. Note that the definition $\mathcal{T}_{-j}$ is more compounded compared to that of $\mathcal{T}$. Indeed, $\mathcal{T}_{-j}$ is the probability that an SBS is both interfering with the test user inside $\mathcal{D}(r_n)$ (as was $\mathcal{T}$ defined for SBSs outside $\mathcal{D}(r_n)$) and is not catering file c_j (since it is inside $\mathcal{D}(r_n)$). Conditioning on the distance r_n, the coverage probabilities $\mathcal{C}_n$ and $\mathcal{C}'_n$ are given by:

$$\begin{aligned}
\mathcal{C}_n(r_n) &= e^{\frac{-\beta r_n^{\eta}\sigma_n^2}{P}}\ \mathcal{L}_{\mathcal{I}_{out}}\left(\frac{\beta r_n^{\eta}}{P}\right), \\
\mathcal{C}'_n(r_n) &= e^{\frac{-\beta r_n^{\eta}\sigma_n^2}{P}}\ \mathcal{L}_{\mathcal{I}_{out}}\left(\frac{\beta r_n^{\eta}}{P}\right)\ \mathcal{L}_{\mathcal{I}_{in}}\left(\frac{\beta r_n^{\eta}}{P}\right),
\end{aligned} \tag{3.15}$$

where $\mathcal{L}_{\mathcal{I}}(t) = \mathbb{E}[e^{-t\mathcal{I}}]$ denotes the Laplace transform (LT) of $\mathcal{I}$, and thus $\mathcal{L}_{\mathcal{I}_{in}}$ and $\mathcal{L}_{\mathcal{I}_{out}}$ denote the LTs of $\mathcal{I}_{in}$ and $\mathcal{I}_{out}$, respectively. Exploiting Approximation 3.1

$$\mathcal{C}_n = \frac{2(\pi\lambda_b)^n}{\Gamma(n)} \int_0^\infty v^{2n-1} e^{\frac{-\beta v^\eta \sigma_n^2}{P}} \exp\Bigg(- \frac{2\pi\mathcal{T}\lambda_b\beta v^2}{\eta-2} \, {}_2F_1\left(1, 1-\frac{2}{\eta}; 2-\frac{2}{\eta}; -\beta\right)\Bigg) e^{-\pi\lambda_b v^2} dv,$$

$$\mathcal{C}'_n = \mathcal{C}_n \left[(1 - \mathcal{T}_{-j} + 2\mathcal{T}_{-j}\vartheta(\beta,\eta))^{n-1} - (1-\mathcal{T}_{-j})^{n-1} \right]. \tag{3.18}$$

---

and following [74, Lemma 1] and [75, Appendix A], the LTs of the aggregate interference from SBSs outside and inside $\mathcal{D}(r_n)$ can be obtained as:

$$\mathcal{L}_{\mathcal{I}_{out}}\left(\frac{\beta r_n^\eta}{P}\right) = \exp\left(-\frac{2\pi\mathcal{T}\lambda_b\beta r_n^2}{\eta-2} \, {}_2F_1\left(1, 1-\frac{2}{\eta}; 2-\frac{2}{\eta}; -\beta\right)\right) \tag{3.16}$$

and

$$\mathcal{L}_{\mathcal{I}_{in}}\left(\frac{\beta r_n^\eta}{P}\right) = \left(1 - \mathcal{T}_{-j} + 2\mathcal{T}_{-j}\vartheta(\beta,\eta)\right)^{n-1} - (1-\mathcal{T}_{-j})^{n-1}$$

where $\quad \vartheta(\beta,\eta) = \beta^{\frac{2}{\eta}} \int_{z=0}^{\beta^{-\frac{1}{\eta}}} \frac{z}{1+z^{-\eta}} dz \tag{3.17}$

where ${}_2F_1(.,.;.;.)$ is the Gauss hypergeometric function [76]. Note that (3.16) and (3.17) account for the fact that the file requested by the test user (i.e., c_j) is not served within $\mathcal{D}(r_n)$. Consequently, the intensities of the interfering SBSs outside and inside $\mathcal{D}(r_n)$ are given by $\mathcal{T}\lambda_b$ and $\mathcal{T}_{-j}\lambda_b$, respectively. By deconditioning over the PDF of the serving distance $f_{r_n}(r)$ given in (2.12), the absolute coverage probabilities can be expressed by (3.18).

In the special case of interference-limited networks (i.e., $\sigma_n^2 = 0$) and path loss exponent $\eta = 4$ (which is common for urban outdoor environments), the coverage

probabilities in (3.18) reduce to the following closed-form expressions

$$\mathcal{C}_n = \left(1 + \mathcal{T}\sqrt{\beta}\ \arctan(\sqrt{\beta})\right)^{-n},$$
$$\mathcal{C}'_n = \mathcal{C}_n \left[\left(1 - \mathcal{T}_{-j}\sqrt{\beta}\ \arctan(\frac{1}{\sqrt{\beta}})\right)^{n-1} - (1 - \mathcal{T}_{-j})^{n-1}\right]. \tag{3.19}$$

The probabilities $\mathcal{T}$ and $\mathcal{T}_{-j}$ in (3.18) and (3.19) depend on the employed transmission schemes, and are derived in the sequel.

3.4.2 $\mathcal{T}$ and $\mathcal{T}_{-j}$ in Unicast Mode

Exploiting Approximation 3.1, the probability that a generic SBS outside $\mathcal{D}(r_n)$ uses the same channel as the test user $\mathcal{T}$ is characterized in the following lemma

Lemma 3.2 *The probability $\mathcal{T}$ that a generic SBS outside $\mathcal{D}(r_n)$ selects a channel randomly and uniformly out of the $|\mathbf{S}|$ available channels is given by:*

$$\mathcal{T} = 1 - \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathbb{P}\left[N = \kappa\right] \frac{|\mathbf{S}| - \kappa}{|\mathbf{S}|}, \tag{3.20}$$

where $\mathbb{P}[N = \kappa]$ is the PMF of the number of channels used by a generic SBS outside $\mathcal{D}(r_n)$, and is expressed as:

$$\mathbb{P}[N = k] = \sum_{\mathbf{x}\in\mathbf{X}} p_{\mathbf{x}} \sum_{\kappa_a=0}^{\min(k,M)} \sum_{\{\mathbf{A}\subseteq\mathbf{x}:\ |\mathbf{A}|=\kappa_a\}} \left\{\mathbb{P}\left[\sum_{i\in\mathbf{A}} U_i = k\right] \prod_{m\in\mathbf{x}\setminus\mathbf{A}} \mathbb{P}[U_m = 0]\right\}$$
$$, 0 \le k \le |\mathbf{S}|. \tag{3.21}$$

Proof. In the *unicast mode*, given that the number of used channels by a generic SBS outside $\mathcal{D}(r_n)$ to serve its associated users is $N = \kappa$. The probability that the SBS randomly chooses a particular channel is given by

$$\mathcal{T}_{|\kappa} = \begin{cases} \frac{\binom{|\mathbf{S}|-1}{\kappa-1}}{\binom{|\mathbf{S}|}{\kappa}} = \frac{\kappa}{|\mathbf{S}|}, & \text{if } \kappa \leq |\mathbf{S}|. \\ 1, & \text{if } \kappa \geq |\mathbf{S}|. \end{cases} \tag{3.22}$$

Therefore, the unconditional probability can be obtained as

$$\begin{aligned} \mathcal{T} &= \sum_{\kappa=0}^{\infty} \mathcal{T}_{|\kappa} \mathbb{P}[N = \kappa] \\ &= \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathbb{P}[N = \kappa] \frac{\kappa}{|\mathbf{S}|} + \sum_{\kappa=|\mathbf{S}|}^{\infty} \mathbb{P}[N = \kappa] \end{aligned} \tag{3.23}$$

Knowing that $\sum_{\kappa=0}^{\infty} \mathbb{P}[N = \kappa] = 1$, therefore $\sum_{\kappa=|\mathbf{S}|}^{\infty} \mathbb{P}[N = \kappa] = 1 - \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathbb{P}[N = \kappa]$. Thus, by substituting into (3.23), lemma 3.2 is obtained.

∎

On the other hand, $\mathcal{T}_{-j}$ accounts for the SBSs inside $\mathcal{D}(r_n)$ that use the same channel as the test user. Therefore, it can be obtained by (3.20) by replacing $\mathbb{P}[N = \kappa]$ by $\mathbb{P}[N_{-j} = \kappa]$ defined in (3.3), to consider the fact that the requested file c_j is not stored in any of the SBSs within $\mathcal{D}(r_n)$.

3.4.3 $\mathcal{T}$ and $\mathcal{T}_{-j}$ in Multicast Mode

Based on the fact that the number of used channels per SBS is upper bounded by M, and exploiting Approximation 3.1, $\mathcal{T}$ can be given by the following lemma.

Lemma 3.3 *The probability $\mathcal{T}$ that a generic SBS outside $\mathcal{D}(r_n)$ selects a channel randomly and uniformly out of $|\mathbf{S}|$ is given by:*

$$\mathcal{T} = \begin{cases} \sum_{\kappa=0}^{M} \mathbb{P}[N=\kappa]\frac{\kappa}{|\mathbf{S}|}, & \text{if } M < |\mathbf{S}|. \\ 1 - \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathbb{P}[N=\kappa]\frac{|\mathbf{S}|-\kappa}{|\mathbf{S}|}, & \text{if } M \geq |\mathbf{S}|. \end{cases} \tag{3.24}$$

Proof. Similarly, in *the multicast mode*, given that the number of used channels by a generic SBS to serve its associated users, which is bounded by M, is $N = \kappa$. The probability that the SBS randomly chooses a particular channel is given by

$$\mathcal{T}_{|\kappa} = \begin{cases} \frac{\binom{|\mathbf{S}|-1}{\kappa-1}}{\binom{|\mathbf{S}|}{\kappa}} = \frac{\kappa}{|\mathbf{S}|}, & \text{if } \kappa \leq M < |\mathbf{S}|. \\ 1, & \text{if } |\mathbf{S}| \leq \kappa \leq M. \\ 0, & \text{if } \kappa > M. \end{cases} \tag{3.25}$$

Therefore, the unconditional probability can be obtained as

$$\begin{aligned} \mathcal{T} &= \sum_{\kappa=0}^{\infty} \mathcal{T}_{|\kappa} \mathbb{P}[N=\kappa] \\ &= \begin{cases} \sum_{n=0}^{M} \mathbb{P}[N=\kappa]\frac{\kappa}{|\mathbf{S}|} & \text{, if } M < |\mathbf{S}|. \\ \sum_{\kappa=0}^{|\mathbf{S}|-1} \mathbb{P}[N=\kappa]\frac{\kappa}{|\mathbf{S}|} + \sum_{\kappa=|\mathbf{S}|}^{M} \mathbb{P}[N=\kappa] & \\ \quad = 1 - \sum_{\kappa=0}^{|\mathbf{S}|-1} \frac{|\mathbf{S}|-\kappa}{|\mathbf{S}|}\mathbb{P}[N=\kappa] & \text{, if } M \geq |\mathbf{S}|. \end{cases} \end{aligned} \tag{3.26}$$

The last equality comes from the fact that $\sum_{\kappa=0}^{\infty} \mathbb{P}[N=\kappa] = 1$. ∎

Similar to the unicast mode, the $\mathcal{T}_{-j}$ can be given by replacing $\mathbb{P}[N=\kappa]$ in

(3.24) with $\mathbb{P}[N_{-j} = \kappa]$ from (3.10). By substituting (3.18) and (3.19) along with (4.12) into (4.7), we obtain the general and closed-form expressions of the hit probabilities for the considered caching network using both unicast and multicast transmission modes, respectively. Such expressions pave the way for maximizing the hit probabilities through optimizing the caching parameters (b_j and $p_{\mathbf{x}}$) for given system and network parameters ($\lambda_b, \lambda_u, \beta, |\mathbf{S}|$, and a_j), as will be introduced in Section 3.5.

3.5 Optimal Caching Distribution

The optimal caching distribution $\mathbf{P}^* = \{p_{\mathbf{x}}{}^*, \mathbf{x} \in \mathbf{X}\}$ that maximizes the hit probability can be obtained via solving the following formulation:

$$\begin{aligned} \max_{\mathbf{P}} \quad & \mathcal{H}(p_1, p_2, \ldots, p_{\mathbf{x}}) \\ \text{subject to} \quad & 0 \leq p_{\mathbf{x}} \leq 1 \quad , \mathbf{x} = 1, 2, \ldots, |\mathbf{X}| \\ & \sum_{\mathbf{x}=1}^{|\mathbf{X}|} p_{\mathbf{x}} = 1 \end{aligned} \tag{3.27}$$

However, the hit probability depends on the caching distribution $p_{\mathbf{x}}$, $\mathbf{x} \in \mathbf{X}$ in a complex manner. Particularly, the recursive nature of some expressions makes it impossible to analytically characterize (e.g., show convexity) the objective function or find tractable expressions for the optimal solution. Hence, we develop a simple, yet accurate, average approximation for the objective function (i.e., hit probability) assuming that each BS uses exactly $\lfloor \frac{\lambda_u}{\lambda_b} \rfloor$ channels for the unicast

scheme and $\min\left(\lfloor\frac{\lambda_u}{\lambda_b}\rfloor, M\right)$ channels for the multicast scheme. Note that the number of used channels per BS is a function of the optimal caching distribution $p^*_{\mathbf{x}}$ as well as the files popularity. Hence, it is hard to find an explicit expression even for the average number of channels used per BS. To maintain the tractability of the optimization problem, the average number of users per BS $\lfloor\frac{\lambda_u}{\lambda_b}\rfloor$ is used as an overestimate on the average number of channels per BS. For the multicast scheme, the truncation $\min(\lfloor\frac{\lambda_u}{\lambda_b}\rfloor, M)$ is used to ensure that the maximum number of channels used per BS is M. Using such constant overestimates within the optimization problem would lead to a suboptimal solution for p^*_x, nevertheless, the results in Section VI show that the proposed suboptimal caching scheme provides tangible gains when compared to the conventional Zipf and uniform caching strategies. The aforementioned approximation decouples the optimization variables from the recursive equations within the hit probability and makes the objective function concave. In addition, we assume that the network is interference-limited network in which the interference dominates the noise power (i.e., $\sigma^2 = 0$). Under such approximate scenario, we find low-complex suboptimal caching distribution for the unicast and multicast schemes.

3.5.1 Unicast Mode

Consider the average approximation, the number of used channel by a generic SBS is $\lfloor\frac{\lambda_u}{\lambda_b}\rfloor$. Hence, the PMFs $\mathbb{P}[N_{-j} = \kappa]$ and $\mathbb{P}[N = \kappa]$ in (3.3) and (3.21) are

reduced to:

$$\mathbb{P}[\tilde{N}_{-j}=\kappa]=\mathbb{P}[\tilde{N}=\kappa]=\begin{cases}1 & \text{if} \quad \kappa=\lfloor\frac{\lambda_u}{\lambda_b}\rfloor \\ 0 & \text{otherwise}\end{cases}. \tag{3.28}$$

Consequently, the thinning factors $\mathcal{T}$ and $\mathcal{T}_{-j}$ become:

$$\tilde{\mathcal{T}}=\tilde{\mathcal{T}}_{-j}=\frac{\lfloor\lambda_u/\lambda_b\rfloor}{|\mathbf{S}|} \tag{3.29}$$

With the interference-limited network assumption ($\sigma^2=0$), the coverage probability $\mathcal{C}_n$ in (3.18) is simplified to:

$$\tilde{\mathcal{C}}_n=\frac{2(\pi\lambda_b)^n}{\Gamma(n)}\int_0^\infty v^{2n-1}\exp\left(-\pi\lambda_b\left[1+\frac{2\tilde{\mathcal{T}}\beta}{\eta-2}{}_2F_1\left(1,1-\frac{2}{\eta};2-\frac{2}{\eta};-\beta\right)\right]v^2\right)dv. \tag{3.30}$$

The integral $\int_0^\infty v^a\ e^{-\alpha v^2}dv=\frac{\Gamma(\delta)}{2\ \alpha^\delta}$,$\mathrm{Re}(\alpha)>0,\mathrm{Re}(a)>0$ and $\delta=\frac{a+1}{2}$ [76]. Thus, $\tilde{\mathcal{C}}_n$ can be expressed as:

$$\tilde{\mathcal{C}}_n=\left(1+\frac{2\beta\lfloor\frac{\lambda_u}{\lambda_b}\rfloor}{|\mathbf{S}|\,(\eta-2)}{}_2F_1\left(1,1-\frac{2}{\eta};2-\frac{2}{\eta};-\beta\right)\right)^{-n} \tag{3.31}$$

By substituting from (3.29) and (3.31) into (3.18), the $\tilde{\mathcal{C}}_n'$ can be evaluated. Also, the approximate OSA success probability $\mathcal{O}_n$ can be given using (3.28) as follows

$$\tilde{\mathcal{O}}_n=1-\tilde{\mathcal{P}}_{n-1}(|\mathbf{S}|) \tag{3.32}$$

where the PMF $\tilde{\mathcal{P}}_{n-1}(\kappa)$ is given by

$$\tilde{\mathcal{P}}_{n-1}(|\mathbf{S}|) = \sum_{t=\lfloor \frac{\lambda_u}{\lambda_b} \rfloor}^{\min\left((n-2)\lfloor \frac{\lambda_u}{\lambda_b} \rfloor, k\right)} \tilde{\mathcal{P}}_{n-2}(t) \binom{\lfloor \frac{\lambda_u}{\lambda_b} \rfloor}{\lfloor \frac{\lambda_u}{\lambda_b} \rfloor - (|\mathbf{S}| - t)} \left(\frac{t}{|\mathbf{S}|}\right)^{\lfloor \frac{\lambda_u}{\lambda_b} \rfloor - (|\mathbf{S}| - t)} \left(1 - \frac{t}{|\mathbf{S}|}\right)^{|\mathbf{S}| - t}$$

$$\text{where} \quad \tilde{\mathcal{P}}_1(|\mathbf{S}|) = \begin{cases} 1, & \text{if } |\mathbf{S}| = \lfloor \frac{\lambda_u}{\lambda_b} \rfloor \\ 0, & \text{otherwise.} \end{cases} . \tag{3.33}$$

Thus, the hit probability can be expressed as

$$\tilde{\mathcal{H}} = \sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j \left(1 - b_j\right)^{n-1} \psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) \tag{3.34}$$

where $\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) = \tilde{\mathcal{C}}_n \, \tilde{\mathcal{O}}_n + \, \tilde{\mathcal{C}}'_n \left(1 - \tilde{\mathcal{O}}_n\right)$

The approximate hit probability is a function of $p_{\mathbf{x}}$ through b_j, via the linear relation between b_j and $p_{\mathbf{x}}$ in (2.3). Thus, the optimization problem in (3.27) is equivalent to [49, Lemma 2]

$$\begin{aligned} \tilde{\mathcal{H}}^* = & \max_{\mathbf{b}} && (3.34) \\ & \text{subject to} && 0 \leq b_j \leq 1 \quad , j = 1, 2, \ldots, J \\ & && \sum_{j=1}^{J} b_j = M. \end{aligned} \tag{3.35}$$

The concavity of the objective function in (3.35) (i.e., $\tilde{\mathcal{H}}$) can be verified by rewriting the hit probability $\tilde{\mathcal{H}}$ as the complementary of the miss probability, it

becomes a separable function w.r.t $b_j, \forall j \in \mathbf{J}$ as follows [1, Lemma 1]:

$$\tilde{\mathcal{H}} = 1 - \sum_{j=1}^{J} a_j \sum_{n=0}^{\infty} (1-b_j)^n \, \psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) \tag{3.36}$$

Note that $\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|)$ is not depend on the caching distribution b_j. Thus, it is easy to confirm that the first derivative $\frac{\mathrm{d}\tilde{\mathcal{H}}}{\mathrm{d}b_j}$ is non-negative and the second derivative $\frac{\mathrm{d}^2\tilde{\mathcal{H}}}{\mathrm{d}b_j^2}$ is non-positive. Thus, $\tilde{\mathcal{H}}$ is a concave function of $(b_1, b_2, \cdots, b_J)$.

The concavity of $\tilde{\mathcal{H}}$ and the linear constraints in (3.35) imply that the Karush-Kuhn-Tucker (KKT) conditions provide necessary and sufficient conditions for optimality. Therefore, the Lagrangian function of (3.35) is given by:

$$L(\mathbf{b}, \mathbf{w}, \mu, v) = \sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j (1-b_j)^{n-1} \psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) + v \left(M - \sum_{j=1}^{J} b_j \right) + \sum_{j=1}^{J} w_j (b_j - 1) - \sum_{j=1}^{J} \mu_j \, b_j. \tag{3.37}$$

Therefore, the sub-optimal file placement b_j^* is the solution of the above convex optimization problem, which is given by the following lemma.

Lemma 3.4 *The sub-optimal file caching strategy that maximizes the approximate hit probability in the unicast mode is given by:*

$$b_j^* = \begin{cases} 0 & , v^* < a_j \Big(\psi_1(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) - \psi_2(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) \Big) \\ 1 & , v^* > a_j \sum_{n=1}^{\infty} \psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) \\ \xi(v^*) & , otherwise \end{cases}, \tag{3.38}$$

where $\xi(v^*)$ *is the solution of* $v^* = a_j\big[\psi_1(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) + \sum_{n=2}^{\infty}(1 - nb_j^*)(1 - b_j^*)^{n-2}\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|)\big]$ *that satisfies* $\sum_{j=1}^{J} b_j = M$.

Proof. Let $\mathbf{b}^*, \mathbf{w}^*, \mu^*$ and v^* be the primal and dual optimal. The KKT conditions for the optimization problem in (3.35) are given by

$$\sum_{j=1}^{J} b_j^* = M \tag{3.39}$$

$$0 \leq b_j^* \leq 1 \quad , j = 1, 2, \ldots, J \tag{3.40}$$

$$\mu_j^*\ b_j^* = 0 \tag{3.41}$$

$$w_j^*(b_j^* - 1) = 0 \quad , j = 1, 2, \ldots, J \tag{3.42}$$

$$\begin{aligned} w_j^* - \mu_j^* - v^* + a_j \Bigg[& \sum_{n=1}^{\infty}(1 - b_j^*)^{n-1}\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|) \\ & - b_j^* \sum_{n=2}^{\infty}(n-1)(1 - b_j^*)^{n-2}\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|)\Bigg] = 0 \end{aligned} \tag{3.43}$$

We write $\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|)$ as ψ_n for simplification, therefore

$$w_j^* = \mu_j^* - a_j\Big[\psi_1 + \sum_{n=2}^{\infty}(1 - nb_j^*)(1 - b_j^*)^{n-2}\psi_n\Big] + v^* \tag{3.44}$$

From (3.41), (3.42) and (3.43), we have

$$w_j^* = -b_j^*\Big[a_j\psi_1 + a_j\sum_{n=2}^{\infty}(1 - nb_j^*)(1 - b_j^*)^{n-2}\psi_n - v^*\Big] \tag{3.45}$$

which, when inserted into (3.42), gives

$$-b_j^*(b_j^* - 1)\Big[a_j\psi_1 + a_j\sum_{n=2}^{\infty}(1 - nb_j^*)(1 - b_j^*)^{n-2}\psi_n - v^*\Big] = 0 \tag{3.46}$$

From (3.46), $0 < b_j^* < 1$ only if

$$v^* = a_j \Big[\psi_1 + \sum_{n=2}^{\infty} (1 - n b_j^*)(1 - b_j^*)^{n-2} \psi_n \Big] \tag{3.47}$$

Since $0 \leq b_j^* \leq 1$, this implies that v^* is bounded by

$$v^* \in \left[a_j \left(\psi_1 - \psi_2 \right), a_j \sum_{n=1}^{\infty} \psi_n \right] \tag{3.48}$$

If $v^* < a_j \left(\psi_1 - \psi_2 \right)$, for small positive ϵ, we have

$$\mu_j^* = w_j^* + a_j \sum_{n=2}^{\infty} (1 - n b_j^*)(1 - b_j^*)^{n-2} \psi_n + a_j \psi_2 + \epsilon > 0 \tag{3.49}$$

Thus, from (3.41) we obtain $b_j^* = 0$. Similarly, if $v^* > a_j \sum_{n=1}^{\infty} \psi_n$, we have

$$w_j^* = \mu_j^* - a_j \sum_{n=2}^{\infty} (1 - n b_j^*)(1 - b_j^*)^{n-2} \psi_n + a_j \sum_{n=1}^{\infty} \psi_n + \epsilon > 0 \tag{3.50}$$

Therefore, from (3.42) we find $b_j^* = 1$. Finally, the sub-optimal file placement b_j^* in (3.38) is obtained. ∎

The sub-optimal caching $b_j^*, \forall j \in \mathbf{J}$ in (3.38) is obtained via the bi-section method in a similar way to the one in [1, Algorithm]. Finally, the sub-optimal per combination caching $\mathbf{P}^* = (p_{\mathbf{x}}^*)_{\mathbf{x} \in \mathbf{X}}$ is obtained by solving the following feasibility

problem [77, Section 4.1]

$$\begin{aligned} \text{Find} \quad & \mathbf{P}^* = (p^*_{\mathbf{x}})_{\mathbf{x}\in\mathbf{X}} \\ \text{subject to} \quad & 0 \leq p^*_{\mathbf{x}} \leq 1 \quad , \mathbf{x} = 1, 2, \ldots, |\mathbf{X}| \\ & \sum_{\mathbf{x}\in\mathbf{X}_j} p^*_{\mathbf{x}} = b^*_j \quad , j = 1, 2, \ldots, J \\ & \sum_{\mathbf{x}=1}^{|\mathbf{X}|} p^*_{\mathbf{x}} = 1 \end{aligned} \tag{3.51}$$

We solve the linear programming in (3.51) using the simplex method. The complexity of (3.51) is highly reduced by eliminating the set of combinations that have zero caching probabilities. More precisely, given the sub-optimal solution of problem (3.35) $b^*_j, \forall j \in \mathbf{J}$, the caching distribution $p_{\{\mathbf{x}\in\mathbf{X}_j : b^*_j=0\}} = 0$ and $p_{\{\mathbf{x}\in\mathbf{X}_{-j} : b^*_j=1\}} = 0$.

3.5.2 Multicast Mode

Similar to the unicast mode, the average approximation assumption in the multicast mode implies that the number of channels used per SBS is $\tilde{M} \equiv \min\left(\lfloor\frac{\lambda_u}{\lambda_b}\rfloor, M\right)$. Therefore, $\mathbb{P}[N_{-j} = \kappa]$ and $\mathbb{P}[N = \kappa]$ become:

$$\mathbb{P}[\tilde{N}_{-j} = \kappa] = \mathbb{P}[\tilde{N} = \kappa] = \begin{cases} 1 & , \kappa = \tilde{M} \\ & \\ 0 & , \text{otherwise} \end{cases} . \tag{3.52}$$

Consequently, the thinning factors $\mathcal{T}$ and $\mathcal{T}_{-j}$ are reduced to:

$$\tilde{\mathcal{T}} = \tilde{\mathcal{T}}_{-j} = \begin{cases} \frac{\tilde{M}}{|\mathbf{S}|} & , \tilde{M} < |\mathbf{S}| \\ 1 & , \tilde{M} \geq |\mathbf{S}| \end{cases}. \tag{3.53}$$

With $\sigma^2 = 0$, the coverage probability $\tilde{\mathcal{C}}_n$ in (3.18) is simplified as:

$$\tilde{\mathcal{C}}_n = \left(1 + \frac{2\beta\tilde{\mathcal{T}}}{\eta - 2} {}_2F_1\left(1, 1 - \frac{2}{\eta}; 2 - \frac{2}{\eta}; -\beta\right)\right)^{-n}. \tag{3.54}$$

Given that each SBS uses $\tilde{M}$ channels, the OSA success probability $\tilde{\mathcal{O}}_n$ is given by:

$$\begin{aligned} \tilde{\mathcal{O}}_n &= 1 - \tilde{\mathcal{P}}_{n-1}\left(|\mathbf{S}|\right) \\ &\approx 1 - \sum_{t=\tilde{M}}^{\min\left((n-2)\tilde{M}, |\mathbf{S}|\right)} \tilde{\mathcal{P}}_{n-2}(t) \binom{\tilde{M}}{\tilde{M} - (|\mathbf{S}| - t)} \left(\frac{t}{|\mathbf{S}|}\right)^{\tilde{M} - (|\mathbf{S}| - t)} \left(1 - \frac{t}{|\mathbf{S}|}\right)^{|\mathbf{S}| - t} \\ \text{where} \quad \tilde{\mathcal{P}}_1(|\mathbf{S}|) &= \begin{cases} 1, & \text{if } |\mathbf{S}| = \tilde{M} \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{3.55}$$

Therefore, the approximate hit probability $\tilde{\mathcal{H}}$ is given by

$$\tilde{\mathcal{H}} = \sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j \left(1 - b_j\right)^{n-1} \varphi_n\left(\beta, \tilde{M}, |\mathbf{S}|\right), \tag{3.56}$$

where $\varphi_n\left(\beta, \tilde{M}, |\mathbf{S}|\right) = \tilde{\mathcal{C}}_n\ \tilde{\mathcal{O}}_n + \ \tilde{\mathcal{C}}'_n\left(1 - \tilde{\mathcal{O}}_n\right)$ The approximate hit probability is a function of $p_{\mathbf{x}}$ through b_j, via the linear relation between b_j and $p_{\mathbf{x}}$ in (2.3).

Thus, the optimization problem in (3.27) turns to:

$$\begin{aligned} \tilde{\mathcal{H}}^* = \max_{\mathbf{b}} \quad & \sum_{j=1}^{J} a_j \sum_{n=1}^{\infty} b_j \left(1-b_j\right)^{n-1} \varphi_n\left(\beta, \tilde{M}, |\mathbf{S}|\right) \\ \text{subject to} \quad & 0 \leq b_j \leq 1 \quad , j = 1, 2, \ldots, J \\ & \sum_{j=1}^{J} b_j = M. \end{aligned} \tag{3.57}$$

The optimization problem in (3.57) is similar to (3.35) with $\varphi_n(\beta, \tilde{M}, |\mathbf{S}|)$ substituting $\psi_n(\beta, \frac{\lambda_u}{\lambda_b}, |\mathbf{S}|)$. The objective function in (3.57) is a concave function of $b_j, \forall j \in \mathbf{J}$ similar to in (3.35) because $\varphi_n(\beta, \tilde{M}, |\mathbf{S}|)$ does not depend on b_j. Therefore, following the same procedure, b_j^* can be expressed as:

$$b_j^* = \begin{cases} 0 & , v^* < a_j \left(\varphi_1\left(\beta, \tilde{M}, |\mathbf{S}|\right) - \varphi_2\left(\beta, \tilde{M}, |\mathbf{S}|\right)\right) \\ 1 & , v^* > a_j \sum_{n=1}^{\infty} \varphi_n\left(\beta, \tilde{M}, |\mathbf{S}|\right) \\ \xi(v^*) & , \text{otherwise} \end{cases} , \tag{3.58}$$

where $\xi(v^*)$ is the solution of $v^* = a_j\big[\varphi_1(\beta, \tilde{M}, |\mathbf{S}|) + \sum_{n=2}^{\infty}(1 - nb_j^*)(1 - b_j^*)^{n-2}\varphi_n(\beta, \tilde{M}, |\mathbf{S}|)\big]$ that satisfies $\sum_{j=1}^{J} b_j = M$. The sub-optimal caching $\mathbf{P}^* = (p_{\mathbf{x}}^*)_{\mathbf{x}\in\mathbf{X}}$ is obtained by solving the same linear programming problem in (3.51).

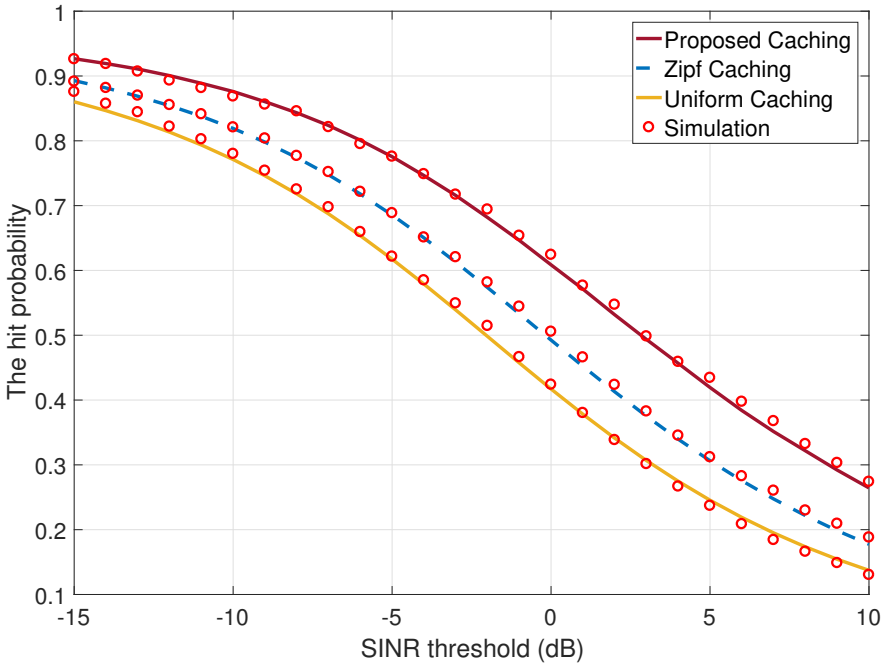
3.6 Simulation & Numerical Results

This section validates the developed mathematical model via Monte Carlo simulations, and discusses several design insights for the considered caching network. In each simulation run of the Monte Carlo simulation, two independent PPPs with intensities λ_b and λ_u are generated in a 10×10 km^2 area. The SBSs independently cache the popular files according to the considered scenario, namely proposed, Zipf or uniform. The user requests are realized using a Zipf distribution with parameter γ. Each users is associated to the nearest SBS that caches the requested file. The SBSs cater the files to the users in unicast and multicast modes using OSA. The simulation is repeated 1000 times and the hit probability is recored for the test user at the origin. Unless otherwise stated, the network parameters are selected as follows; $\lambda_b = 4$ SBSs/Km^2, $P = 1$ watt, and $\eta = 4$.

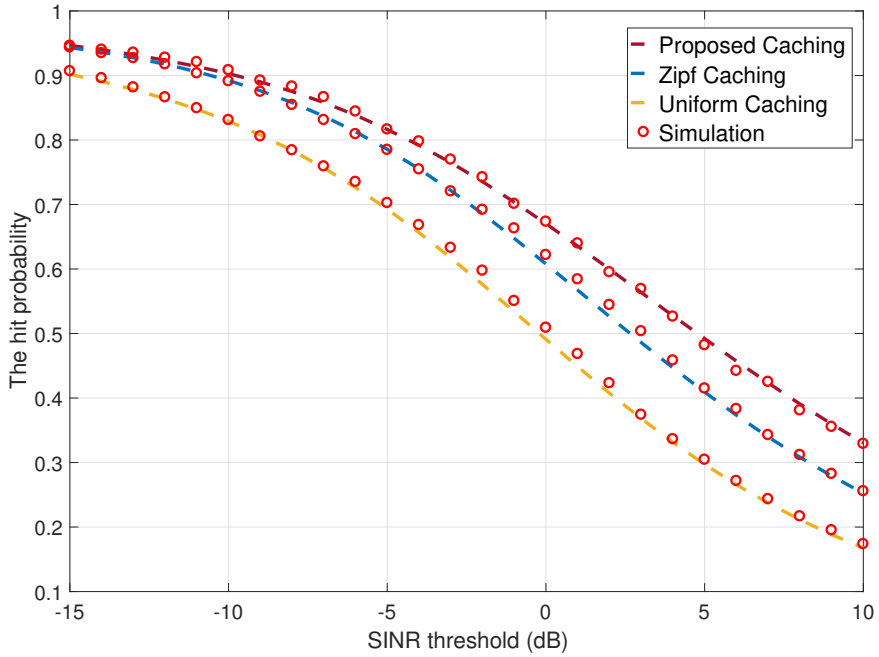
Fig. 3.1 shows the hit probability versus the SINR threshold for both the multicast and unicast transmission modes. The figure shows good matching between the analytical and simulation results, which validates the developed mathematical model. Fig. 3.1 also illuminates the gain in terms of the hit probability when adopting our proposed sub-optimal caching compared to Zipf (i.e., file popularity based) and uniform caching schemes. Apparently, the hit probability gain achieved by employing the proposed caching is more significant for the unicast mode (Fig. 3.1a) compared to the multicast mode (Fig. 3.1b), and also, for the high SINR threshold compared to the low SINR threshold. This observation can be explained by the cost of being catered from a farther SBSs. Particularly, when

the interference is significant (i.e., in the unicast and/or high SINR threshold), the coverage probability from an SBS other than the geographically nearest one becomes too small. Hence, the proposed caching is highly desirable to maximize the probability of finding the desired file in the closest possible SBSs. However, when the interference significance is reduced (via multicast or lower SINR threshold), the cost of fetching the file from a farther SBS decreases and the effect of caching becomes less impactful.

It worth noting that we do not aim to quantitatively compare the performances of the unicast and the multicast schemes. Instead, we want to study the qualitative response and sensitivity of each scheme to the network parameters. However, due to space limitations, we combine the unicast and multicast performances in the rest figures. Fig. 3.2 shows the hit probability performance versus the number of channels. The figure manifests the superiority of the proposed caching over the Zipf and uniform schemes for both transmission modes. The figure also shows that the significance of the proposed caching gain in terms of hit probability decreases when increasing the number of channels. Thanks to the OSA, the interference is relieved as the number of channel increases, which reduces the cost of fetching the file from a farther SBS. That is, increasing the number of channels makes farther SBSs more capable of successfully catering for a file request, which decreases the significance of finding the file in a nearer SBS. Thus, the gain of the proposed caching over the Zipf and uniform caching decreases as the number of channels increases. It can also be observed that proposed caching significant is more no-



(a) The unicast mode ($|S| = 30 \,\&\, J = 15 \,\&\, K = 5 \,\&\, \gamma = 1.8 \,\&\, \frac{\lambda_u}{\lambda_b} = 20$)



(b) The multicast mode ($|S| = 10 \,\&\, J = 20 \,\&\, K = 5 \,\&\, \gamma = 1.8 \,\&\, \frac{\lambda_u}{\lambda_b} = 30$)

Figure 3.1: The hit probability ($\mathcal{H}$) vs. the SINR threshold (β) for both the multicast and unicast modes compared to the Zipf and uniform caching schemes.

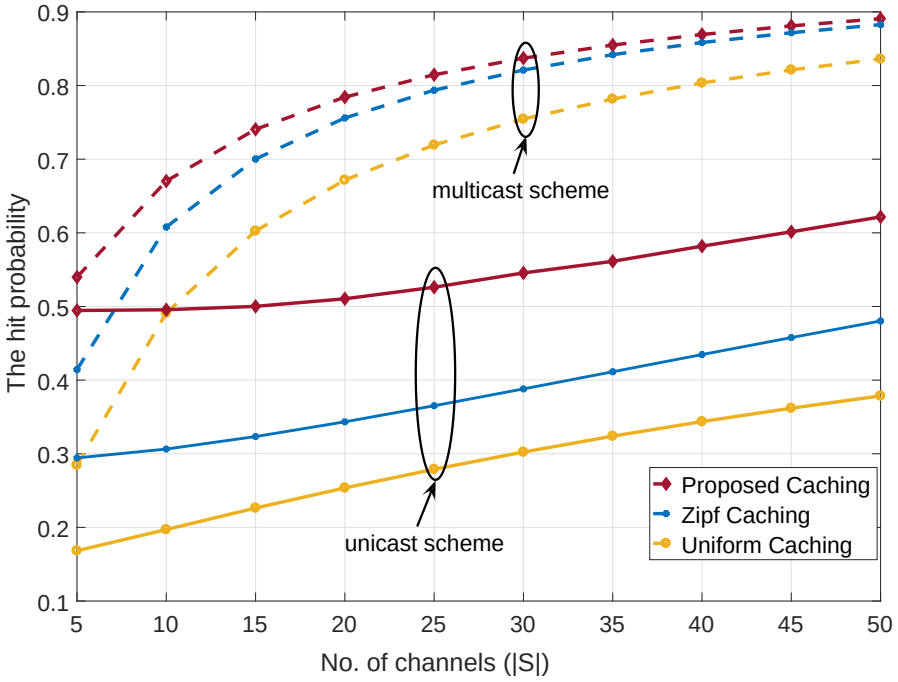


Figure 3.2: $\mathcal{H}$ vs. M ($|\mathbf{S}| = 10 \,\&\, J = 15 \,\&\, \gamma = 1.8 \,\&\, \lambda_u/\lambda_b = 10 \,\&\, \beta = 0$ dB)

ticeable in the unicast mode than in the multicast mode, due to the lower OSA success and higher interference effects in the unicast mode, thus calling more for optimized caching.

The effect of fetching the files from a closer SBS is highlighted in Figs. 3.3 and 3.4 by increasing the cache size and increasing the Zipf parameter, respectively. Fig. 3.3 shows that increasing the cache size improves the hit probability due to the increased probability of finding the required file in a closer BS. This highlights the tradeoff between the caching system performance and the network storage resources. Fig. 3.4 shows that an increasing Zipf parameter improves the hit probability, due to the dominant request of users to less number of files. Hence, this increases the probability of finding the dominant popular files in closer SBSs. In both figures, it is clear that the hit probability gain due to employing the

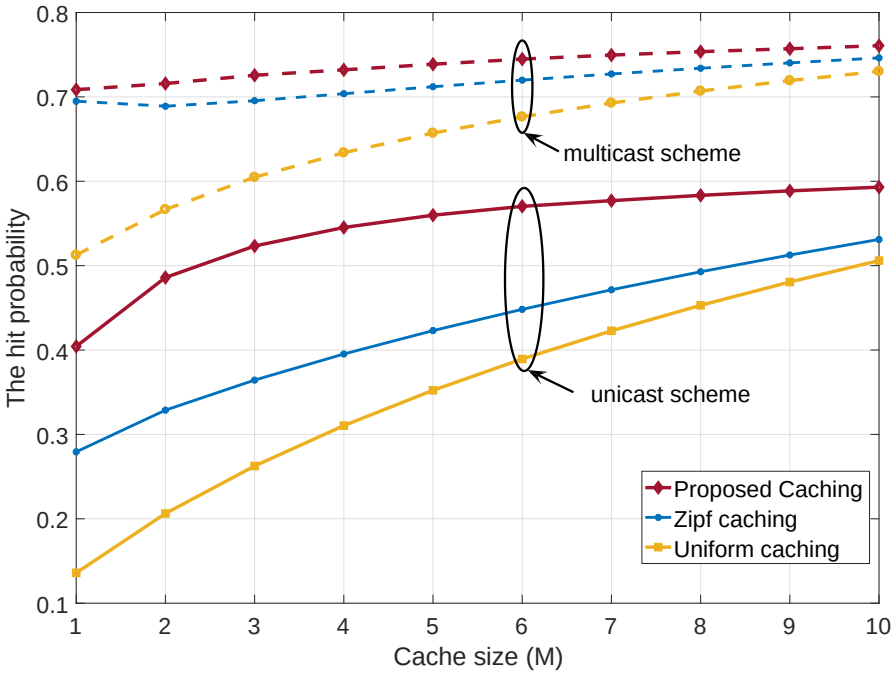


Figure 3.3: $\mathcal{H}$ vs. M ($|\mathbf{S}| = 10 \,\&\, J = 15 \,\&\, M = 5 \,\&\, \gamma = 1.8 \,\&\, \lambda_u/\lambda_b = 10 \,\&\, \beta = 0\,\text{dB}$).

proposed cache strategy is less significant in multicast mode when compared to the unicast mode, which manifests the cost of interference on the hit probability when fetching the file from farther SBSs.

Fig. 3.5 plots the hit probability versus the average number of users per SBS (i.e., λ_u/λ_b). As expected, the figure shows that the hit probability decreases as the average number of users per a SBS increases. It can also be noticed that the unicast mode exhibits a faster decaying trend, particularly at high β compared with the multicast scheme, thanks to decreasing OSA success probability. Indeed, such decrease can be attributed to the higher OSA failure probability at higher β, which increases the interference, and hence, the cost of fetching the requested file from farther SBS.

To better interpret the performance gains between the different caching strate-

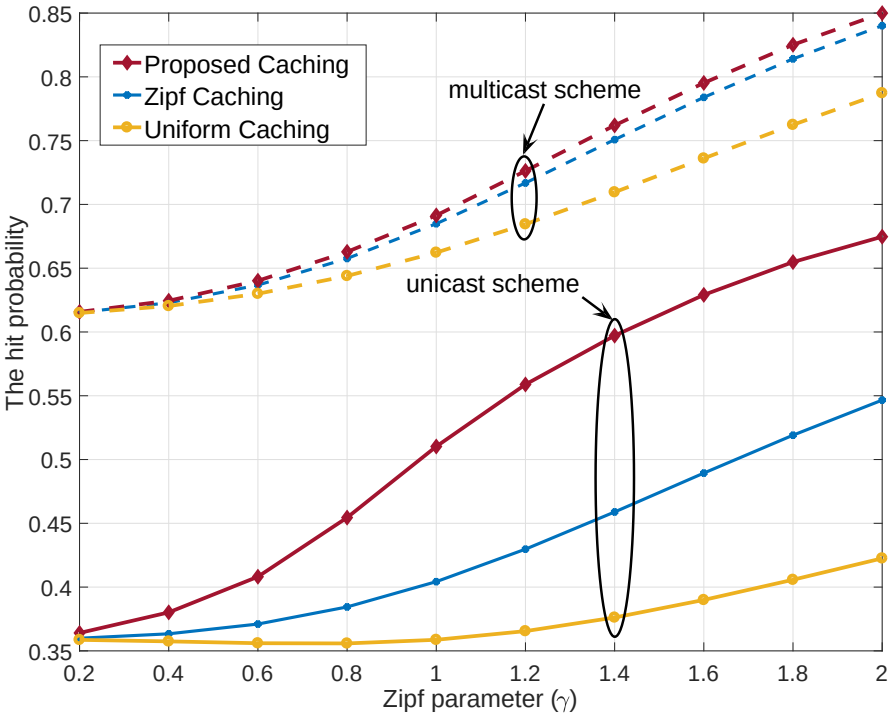


Figure 3.4: $\mathcal{H}$ vs. γ ($|\mathbf{S}| = 10 \,\&\, J = 15 \,\&\, M = 5 \,\&\, \beta = 0$ dB)

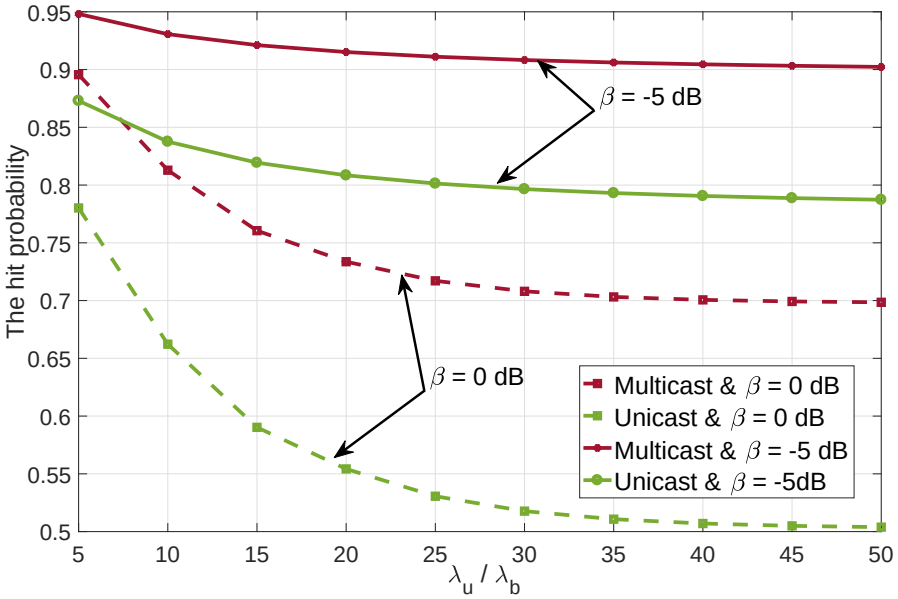
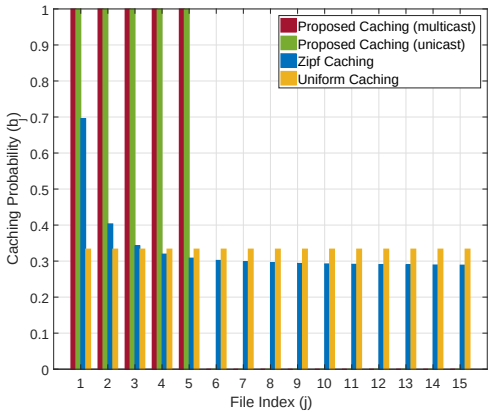
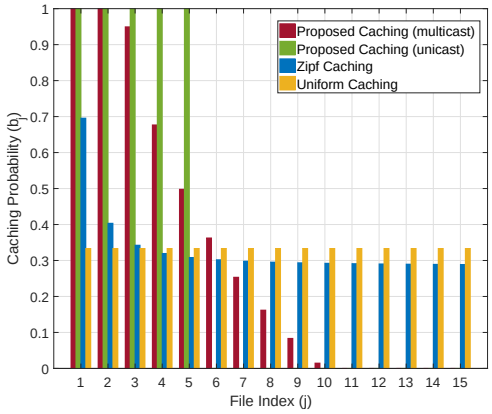


Figure 3.5: $\mathcal{H}^*$ vs. $\frac{\lambda_u}{\lambda_b}$ ($|\mathbf{S}| = 20 \,\&\, J = 15 \,\&\, M = 5 \,\&\, \gamma = 1.8$ dB)

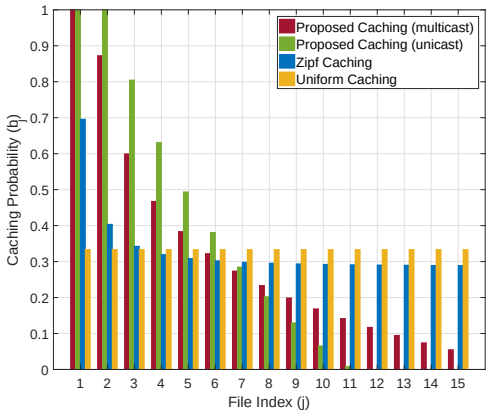
gies, we plot Fig. 3.6 to show the file caching probabilities in the buffer of a generic SBS. First, we note that the probabilities in Fig. 3.6 sum to M, which is the num-



(a) $|\mathbf{S}| = 5 \,\&\, \beta = 0$ dB



(b) $|\mathbf{S}| = 10 \,\&\, \beta = 0$ dB



(c) $|\mathbf{S}| = 15 \,\&\, \beta = -5$ dB

Figure 3.6: The proposed sub-optimal caching b_j^* for both multicast and unicast modes versus the uniform and Zipf distributions ($J = 15 \,\& M = 5 \,\& \gamma = 1.8 \,\& \, \lambda_u/\lambda_b = 10$)

ber of files that a SBS can store. It is obvious that the Zipf caching outperforms the uniform caching as it accounts for the file popularity. The proposed caching further adapts to the network conditions in addition to the file popularity, and hence, it outperforms the Zipf caching. For instance, Fig. 3.6a shows that, when the interference conditions are too adverse (due to the small number of channels and/or high β), the proposed caching strategy caches only the most M popular files in all SBSs for both the multicast and unicast transmission modes. Articulated differently, the proposed caching does not store any of the files c_j for $M < j \leq J$ at any SBS. Indeed, any SBS farther than the closest one fails to cater a requested file with the required SINR threshold β. Consequently, the proposed caching strategy sacrifices the hitting probability of less popular files to guarantee that the most M popular files are always served from the geographically closest SBS. Relieving the interference adversity (i.e., by increasing the number of channels or decreasing β), Figs. 3.6b and 3.6c show that the proposed caching gradually approaches the Zipf distribution, i.e., gradually starts to diversify the caching options over the SBSs according to the files popularity. The figure also shows that the number of sacrificed files (i.e., that are never cached) depends on the cost of fetching the files from farther SBSs.

3.7 Chapter Summary

This chapter develops a mathematical framework, based on stochastic geometry, to characterize the hit probability of a cache-enabled unicast/multicast 5G net-

work with SBS multi-channel capabilities and OSA. Integral forms for the hit probability, which reduces to closed-forms at practical special cases, are obtained. To this end, the sub-optimal caching distribution that maximizes the hit probability is presented. The results manifest the superiority of the multicast transmission mode. The results also signify the importance of OSA and the proposed sub-optimal caching on the hit probability. In particular, at adverse interference conditions, the proposed caching significantly improves the hit probability when compared to the popularity or uniform based caching. When the number of channels increases, the OSA becomes effective and relieves the interference adversity. In this case, the cost of fetching the desired file from a farther SBS decreases and the significance of the caching strategy becomes less distinguishable.

CHAPTER 4

FLEXIBLE CACHING IN MILLIMETER WAVE NETWORKS

4.1 Introduction

Ultra-densification, millimeter wave (mmW) communications, and proactive network-edge caching, utilized within mmW fog networks (mmFNs), are foreseen to provide tangible gains for broadband access, network capacity, and latency. However, caching implementation in mmFN imposes high capital expenditure (CAPEX) due to the ultra-high density of base stations (BSs). For a given caching CAPEX, it may be more efficient to install higher capacity caches in a fraction of the BSs than installing smaller capacity caches in every BSs. In the former case, wireless self-backhauling of mmW systems can be exploited to share the cache

contents stored in a given cache enabled BSs (CE-BSs) with other BSs in the network. In this regards, this chapter develops a mathematical model, based on stochastic geometry, to study the tradeoff between the cache size and intensity of CE-BSs on the probability that requested popular contents are retrieved from the network edge, denoted as the hit probability. Assuming a power-law inverse relationship between the cache size and intensity of CE-BSs, an optimization problem is formulated and solved for the intensity of CE-BSs and probabilistic file placement in caches such that the hit probability is maximized. The results show that neither installing small caches in every BS nor having sufficiently high capacity caches (i.e., that confine all popular files) installed in small number of BSs exploit the full potential of mmFN. Instead, there exists an optimal balance between the cache size and intensity of CE-BSs, which depends on the network parameters such as the applied caching strategy, required rate, total intensity of BSs, popular content distribution, and cache size/intensity relationship.

The 5G cellular networks are challenged to offer multi-fold improvement in broad-band access, network capacity, and end-to-end latency. For instance, augmented/virtual reality applications require data rates in the order of Gbps with latency less than 20ms [78]. To concurrently improve such contradictory key performance indicators (KPIs), a paradigm shift in the network deployment and design is required. In this regards, the evolving millimeter wave (mmW) fog networks (mmFN) offers a promising solution to materialize the 5G vision. Particularly, mmFN brings three novel, and complementary, ingredients to cellular

networks, namely, mmW communications, ultra-densification, and proactive edge-caching [79].

The mmW communications exploit high bandwidth channels within the under-utilized band of 20-100 GHz, which significantly improves transmission rates compared to conventional cellular micro-wave (0.8-2.1 GHz) bands. Ultra-densification mitigates the poor propagation properties of mmW communications (e.g., high path-loss and blockage) by reducing the communication link distances and increasing the probability of line of sight [80]. Furthermore, ultra-densification improves the network capacity though increasing the spatial spectral efficiency and reducing devices load per base station (BS). Proactive caching brings popular contents to the network edge (i.e., nearer to the devices), which reduces communication latency. Hence, the mmFN provides the foreseen low-latency broadband access for dense devices, which is crucial for several 5G applications such as augmented/virtual reality, online gaming, social networking, and video streaming [11].

The potential gains of densification, caching, and mmW communications have attracted tremendous interest from academia and industry. The anticipated performance of ultra-dense networks is characterized in [81]. Experimental proof of concept for mmW communications is conducted in [18, 57]. On the scale of networks, stochastic geometry is utilized in [21, 54–56] to characterize the coverage probability and rate in mmW cellular networks. Self-backhauling in dense mmW cellular networks is studied in [21]. For caching systems, optimal file placement the

maximizes the probability that requested files are retrieved from the BSs' caches, denoted as the hit probability, is derived in [1,36,37]. The performance gain due to opportunistic spectrum access in cache enabled network is characterized in [74,82]. However, none of [1,36,37,74,82] studies caching in mmW networks. The authors in [83, 84] and [23] study caching in mmW networks when, respectively, all BSs and a fraction of BSs are cache enabled. However, none of the above research studies the interactions between the cache size and intensity of cache enabled BSs (CE-BSs) in self-backhauled mmFN (i.e., ultra-dense) networks.

Caching implementation and design for mmFN networks is a challenging task due to the ultra high density of BSs. For a given capital expenditure (CAPEX) to install caches in BSs, there is an obvious inverse relationship between cache size and the intensity of BSs to be augmented with caches. Augmenting all BSs with caches may impose stringent constraints on the cache size. Hence, small portion of the popular files library can be cached at the network edge, and hence, proactive caching is not exploited to its full potential. For the same caching CAPEX, larger cache size can be utilized at the network edge, however, at the cost of implementing the caches at a fraction of the BSs. Note that, exploiting the high density of BSs and self-backhauling of mmW systems, the large caches implemented in CE-BSs can be accessed with cache-free BSs (CF-BSs). The tradeoff between the cache size and density, for the same CAPEX, is a crucial problem that has never been addressed before.

To the best of the authors knowledge, this work is the first to study the cache

size/intensity tradeoff in self-backhauled mmFN for optimized and most popular caching strategies. Particularly, we study a Poisson point process (PPP) mmFN with self-backhauling capabilities where the cache size implemented in the BSs is inversely proportional to the CE-BSs intensity. Without loss of generality, an inverse power-law relationship between the fraction ζ of CE-BSs and the cache capacity $M \propto \frac{1}{\zeta^{\epsilon}}$ at each. Due to the significant advancement in storage devices along with the high cost of casing, N separately cased caches of a certain capacity (e.g., M Gigabytes) is much more expensive than a single cache with the aggregated (i.e., NM Gigabytes) capacity. Hence, the inverse proportionality is non-linear. While any relationship between the cache sizes and intensity of CE-BSs can be incorporated to the developed model, without loss of generality we choose the power-law inverse proportionality with an exponent ($\epsilon \geq 1$).

Due to the poor propagation characteristic of mmW signals, users are always associated to their geographically closest BS. If the geographically closest BS does not store the requested file (i.e., CE-BS that caches a different set of files or CF-BS), then a single-hop wireless backhauling is utilized to retrieve the requested file from the closest CE-BS that stores it. Our results manifest the tradeoff between caching size and intensity of CE-BSs and show that there exists an optimal balance that maximizes the hit probability at the network edge. Such optimal balance depends on the network parameters such as the applied caching strategy, required rate, total intensity of BSs, popular content distribution, and cache size/intensity relationship.

The rest of this chapter is organized as follows: the system model is introduced in Section 4.2. Section 4.3 derives the hit probability expression. In Section 4.3.2, the optimization problem to maximize the hit probability is formulated. Finally, the numerical results are illustrated in Section 4.4.

4.2 System Model

4.2.1 Network Model

We consider a dense mmFN where the BSs are distributed in $\mathbb{R}^2$ according to a PPP $\mathbf{\Psi}$ with intensity λ. To utilize the tradeoff between the cache size and intensity of CE-BSs, caches are not installed in all BSs. Instead, a BS is assumed to have a cache of size $M \propto \frac{1}{\zeta^{\epsilon}}$ with probability $0 < \zeta \leq 1$, where $\epsilon \geq 1$ due to the high cost of cache casing and installation. The cache implementation probability ζ is assumed to be independent across all BSs. Utilizing the independent thinning of the PPP, the CE-BSs constitute a PPP $\mathbf{\Psi}^c \subseteq \mathbf{\Psi}$ with intensity $\zeta\lambda$.

We focus on a mmW downlink network with universal frequency reuse for network access (i.e., BS to UE) and backhauling (i.e., BS to BS) links [21]. All BSs transmit with the same power level of P_t. Without loss of generality, we focus on a UE located at the origin, which becomes the typical user after spatial averaging [62].

4.2.2 Directional Beamforming

The small wavelength of mmW signals enables conveniently sized and highly directive antenna arrays, which can be included into small BSs as well as users equipment (UEs). From the wireless link level, high antenna directivity is critical for compensating the poor propagation properties of mmW signals. At the network level, narrow beamwidth communications mitigate the mutual interference between BSs, which is crucial in ultra-dense deployment [81].

In this work, we assume that both the BSs and UEs are equipped with antenna arrays but with different sizes. The antenna array gains of the BSs and UEs are, respectively, denoted as G_{bs} and G_u, where $G_{bs} > G_u$ due to the more stringent size constraints of UEs. For simplicity, we follow [21, 54–56] and assume a two-state beam pattern for the employed antenna arrays, which is expressed as:

$$G_j(\phi) = \begin{cases} G_j^{\max}, & \text{if } |\phi| \leq \triangle\phi_j, \\ G_j^{\min}, & \text{if } |\phi| > \triangle\phi_j, \end{cases} \tag{4.1}$$

where $j \in \{bs, u\}$, $\phi \in [-\pi, \pi)$ is the deviation angle from the antenna boresight, $\triangle\phi_j$ denotes the beamwidth of the mainlobe, $G_j^{\max}$ and $G_j^{\min}$ are the antenna gains of main and side lobes, respectively. During UE to BS association, perfect antenna alignment between the UE and intended BS is assumed. However, the beams of all interfering links are received at random orientation with respect to each other as well as with respect to the intended receiver.

4.2.3 Propagation Model

For an arbitrary distance r between the transmitting and receiving nodes, the received power is given by $P_t\,\psi\,\ell(r)$ where ψ is the effective directivity gain due to the relative orientations of the transmit and receive antennas, and $\ell(\text{dB})$ is the distance dependent path loss. The path-loss is expressed as $\ell(\text{dB}) = \delta + 10\alpha\log(r) + \vartheta$, where δ is the path loss at a close-in reference distance, α is the path loss exponent, and $\vartheta \sim \mathcal{N}(0, \eta^2)$ is the normally distributed (i.e., log-normal in the absolute scale) shadowing with variance η^2. Motivated by the studies in [21] and [55], which assume each of the access and backhaul links has its own propagation parameters denoted, respectively, via the tuples $(\delta_a, \alpha_a, \eta_a^2)$ and $(\delta_b, \alpha_b, \eta_b^2)$.

4.2.4 Blockage Model

Following the field measurements and stochastic blockage models in [18, 57], we assume that the probability that a link is line of sight (LOS) is $e^{-\rho r}$, where the decaying rate ρ depends on the building parameters and density. For analytical tractability, the LOS probability function can be approximated by step functions, where the irregular geometry of the LOS region is replaced via a LOS ball that would lead to an equivalent distribution for the signal-to-interference-plus-noise-ratio (SINR). It is worth pointing out that such step functions approximation is utilized and validated in [29, 85]. In this paper, we adopt a single-ball model with transition radius of R_l such that all link distances of $r < R_l$ can be LOS

with probability $\mathcal{P}_l$ and NLOS with probability $(1 - \mathcal{P}_l)$. On the other hand, the transmission link of distance $r \geq R_l$ is considered non-LOS (NLOS) with probability one.

4.2.5 Content Popularity and Caching Models

We consider a finite library of popular files (contents), denoted by $\mathbf{J} = \{c_1, c_2, ..., c_J\}$. It is assumed that all files have the same length. However, this analysis can be still applied for files of different sizes by chopping each file into equal length packets. We assume that the files popularity is fully known a priori, via big data analytics or machine learning techniques, for the network operator. The content popularity is assumed to follow the Zipf distribution due to its practical relevance [63]. The files popularity distribution is expressed as:

$$a_j = \frac{j^{-\gamma}}{\sum_{i=1}^{J} i^{-\gamma}}, \tag{4.2}$$

where a_j is the probability that a UE requests the file c_j, and γ is the Zipf parameter that governs the popularity distribution skewness. Larger (smaller) γ increases (decreases) the discrepancies among the files popularity and implies that fewer (more) files are frequently requested. Without loss of generality, it is assumed that the files of the library are enumerated in a descending order of their popularity, i.e., $a_1 \geq a_2 \geq \cdots \geq a_J$.

As discussed before, the cache size is assumed to be inversely proportional to the intensity of CE-BSs $0 < \zeta\lambda \leq \lambda$. The cache size is given as $M = \left\lfloor \frac{M_o}{\zeta^{\epsilon}} \right\rfloor$ files,

where M_o is the proportionality constant (i.e., the smallest cache capacity for a given CAPEX when $\zeta = 1$), and $\lfloor x \rfloor$ is the floor operator. Note that the extreme case of $\zeta = 0$ is eliminated from this paper because it depicts a scenario where no caches are installed in the network, which is out of the scope of this paper. Instead, we assume that the minimum fraction of CE-BSs ζ corresponds to the case where the capacity of each installed cache is large enough to store all the popular J-files. Consequently, $(\frac{M_o}{J})^{1/\epsilon} \leq \zeta \leq 1$ and $M_o \leq M \leq J$.

We adopt a probabilistic file placement technique, where the probability that a CE-BS stores the file c_j is $b_j (0 \leq b_j \leq 1), \forall c_j \in \mathbf{J}$. To avoid duplicate file caching, we adopt the probabilistic caching strategy proposed in [1]. For any tuple (M, J, b_j), M equal rows of unit size are sequentially filled with the probabilities b_j. If the full capacity of a row is reached before encompassing a given b_j, the remaining portion of that b_j is continuously filled in the next row. Finally, a random number within $[0, 1]$ is selected to determine the selected file combination. Since $\sum_{j=1}^{J} b_j = M$, all popular files are considered within such caching strategy. Fig. 4.1 illustrates an example of the adopted probabilistic content placement strategy for $M = 4$ and $J = 10$, where a vertical line uniformly placed in the region [0,1] is used to choose the $M \leq J$ files to be stored in the CE-BS. It is clear from the illustrative example that the event of storing a file c_j occurs according to the probability b_j and that file duplication is alleviated.

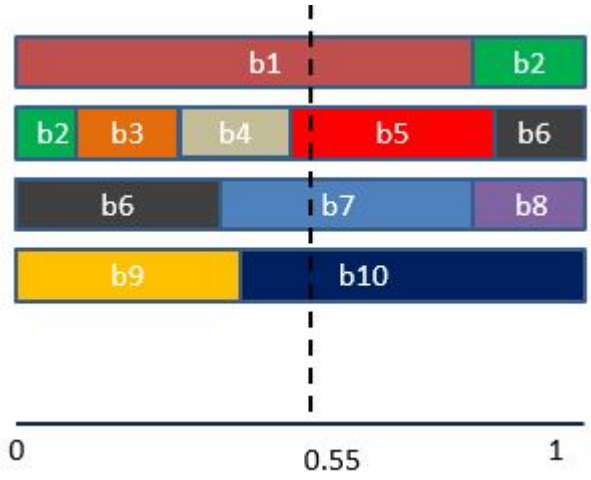


Figure 4.1: An illustration of the probabilistic caching method [1] with $M = 4$ and $J = 10$. The set of stored files is $\{c_1, c_5, c_7, c_{10}\}$

4.2.6 Association model and self backhauling

Due to the poor propagation properties of mmW signals, each UE is associated to its geographically closest BS irrespective of its cache status (i.e., CE-BS or CF-BS). Let x_a^* denote the serving BS for a UE that requests the popular file c_j, where the subscript a denotes the access link. Let the thinned PPP $\mathbf{\Psi}_j^c \subseteq \mathbf{\Psi}^c$ of intensity $\lambda\zeta b_j$ denote the set of CE-BSs that store c_j. Based on the proposed system model, the UE requesting c_j is served with one of the following alternatives:

(i) If $x_a^* \in \mathbf{\Psi}_j^c$ is CE-BS and stores the file c_j, then the requested file is served to the UE from the service BS cache. This event occurs with probability ζb_j.

(ii) The serving BS does not store the file c_j if i) x_a^* is CF-BS, which occurs with probability $(1-\zeta)$; or ii) x_a^* is a CE-BS but does not store c_j, which occurs with probability $\zeta(1-b_j)$. In either of these cases, the requested file is retrieved via the wireless backhaul from the CE-BS that stores c_j and

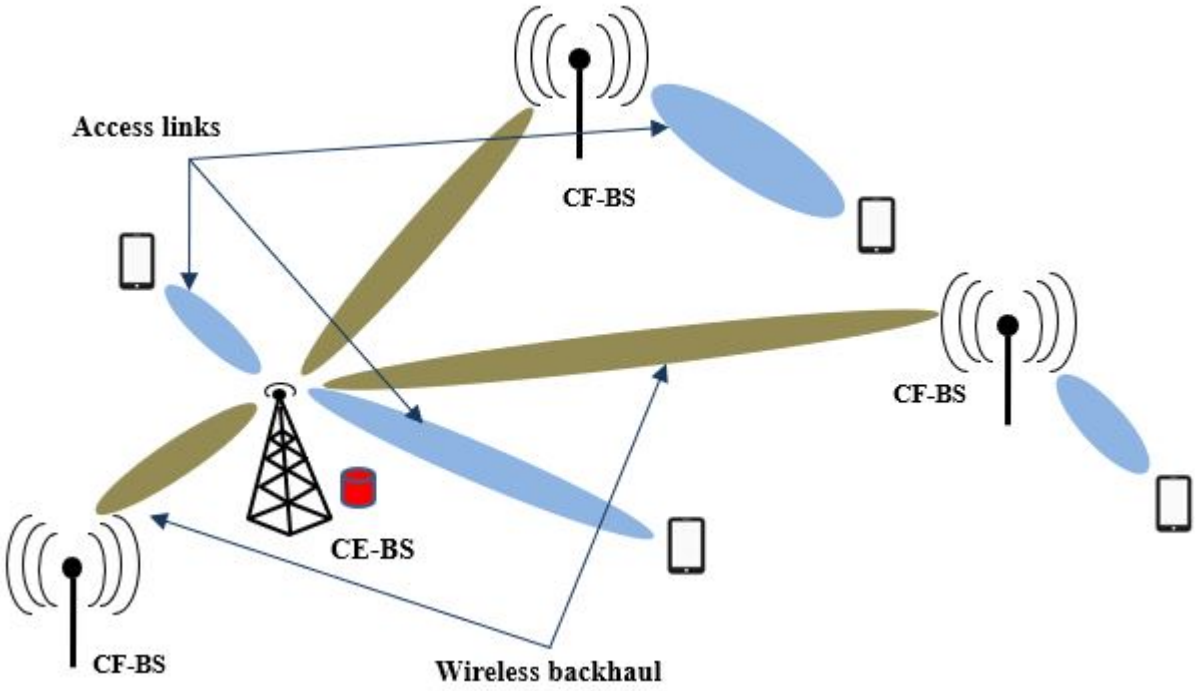


Figure 4.2: Self-backhauled network with the CE-BS providing wireless backhaul to the tagged BSs and access link to the tagged users

is geographically closest to x_a^*. Such backhauling CE-BS is denoted as x_b^*. Hence, when the serving BS does not store the requested file, a two hop transmission (i.e., backhaul from x_b^* to x_a^* then access from x_a^* to the UE) is utilized to serve the UE request.

Fig. 4.2 illustrates the association and self-backhauling policies employed in this paper.

4.2.7 Performance Metric

The hit probability, defined as the probability that the typical UE successfully downloads the requested file from the network edge (i.e., through the CE-BS in one or two hops), is the considered performance metric. Note that a successful

file transmission requires that the received SINR to be above a certain threshold β. Based on the two aforementioned alternatives (i.e., single and two hops) for serving the typical UE, the hit probability can be expressed as

$$\begin{aligned}\mathcal{H} &= \sum_{j=1}^{J} a_j \Big(\zeta b_j \mathbb{P}[\mathrm{SINR}_a > \beta] + (1 - \zeta b_j) \mathbb{P}[\mathrm{SINR}_b > \beta,\, \mathrm{SINR}_a > \beta] \Big), \\ &\overset{(a)}{=} \sum_{j=1}^{J} a_j \left(\zeta b_j S_a(\beta) + (1 - \zeta b_j) S_a(\beta) S_b(\beta) \right), \end{aligned} \tag{4.3}$$

where SINR_a and SINR_b correspond to the SINR of the access and backhaul links, respectively; $S_a(\beta) = \mathbb{P}[\mathrm{SINR}_a > \beta]$ and $S_b(\beta) = \mathbb{P}[\mathrm{SINR}_b > \beta]$ are the successful transmissions probabilities, denoted hereafter as the coverage probabilities, of the access and backhaul links, respectively. The equality (a) is obtained by the fact that $S_a(\beta)$ and $S_b(\beta)$ are independent events.

4.3 The hit probability analysis

This section characterizes the hit probability for the self-backhauled mmFN, where the tranmission success probabilities $S_a(\beta)$ and $S_b(\beta)$ are derived in Section 4.3.1. Then, the cache size and file placement probabilities b_j are optimized in Section 4.3.2.

4.3.1 SNR distribution

Thanks to the narrow-beam mmW transmissions, the mmFN is most likely noise-limited [21, 55, 56, 80]. Therefore, the SINR in (4.13) is replaced with the signal-to-noise-ratio (SNR). To characterize the SNR distribution, we define the one dimensional point process seen by the typical UE for the access network. Such 1D point process is constructed by mapping $\mathbf{\Psi}$ to $\mathbb{R}$, where the mapping function is the access path loss function. Hence, the access 1D point process is defined as $\mathcal{N}_a = \{\ell_a(x) = \frac{|x|^\alpha}{s}\}$, where $s = \{e^{-0.1\delta \ln 10 - 0.1\vartheta \ln 10}\}_{x\in\mathbf{\Psi}}$ which is log normal distributed random variable, i.e., $s \sim \ln\mathcal{N}(m, \sigma)$ with mean $m = -0.1\delta \ln 10$ and variance $0.1\eta \ln 10$. Using the mapping and displacement theorems [71, Section 2.7], the intensity measure for the point process $\mathcal{N}_a$, denoted by $\Lambda_a([0, t))$ is characterized by the following lemma.

Lemma 4.1 *The distribution of the path loss of the access link between the typical UE and its serving BS x_a^*, defined as $\mathbb{P}\left[\ell_a\left(x^*\right) > t\right] = \exp\left(-\Lambda_a\left([0, t)\right)\right)$, is obtained using the avoid property of the PPP process. The intensity measure $\Lambda_a\left([0, t)\right)$ is given by (4.7) at the top of next page, with $m_j = -0.1\delta_j \ln 10$ and $\sigma_j = 0.1\eta_j \ln 10$, with $j \equiv l$ for LOS and $j \equiv n$ for NLOS link and $Q(.)$ is the Q-function.*

Proof. Recall the point process $\{\mathcal{N}_a = \ell_a(x) = \frac{|x|^\alpha}{s}\}_{x\in\mathbf{\Psi}}$ on $\mathbb{R}$ formed by the path loss from each BS to the typical user at the origin. Using the displacement

theorem, the intensity measure $\Lambda_a((0,t])$ can be expressed by

$$\begin{aligned}\Lambda_a([0,t)) &= \int_{\mathbb{R}^2} \mathbb{P}[\ell(x) < t]\mathrm{d}x = 2\pi\lambda \int_{\mathbb{R}} \mathbb{P}\left[\frac{r^{\alpha(r)}}{s(r)} < t\right] r\mathrm{d}r \\ &= 2\pi\lambda\Bigg\{ \sum_{j\in\{l,n\}} \mathcal{P}_j \int_{\mathbb{R}} \mathbb{P}\left[r < (ts_j)^{1/\alpha_j}\right] \mathbb{1}_{(r<R_l)}\, r\mathrm{d}r \\ &\qquad + \int_{\mathbb{R}} \mathbb{P}\left[r < (ts_n)^{1/\alpha_n}\right] \mathbb{1}_{(r\geq R_l)}\, r\mathrm{d}r \Bigg\} \end{aligned} \tag{4.4}$$

where $\mathcal{P}_j = \{\mathcal{P}_l, 1-\mathcal{P}_l\}$ for $j \in \{l,n\}$. Therefore,

$$\begin{aligned}\Lambda_a([0,t)) &= 2\pi\lambda\, \mathbb{E}_s\Bigg\{ \sum_{j\in\{l,n\}} \mathcal{P}_j \Bigg[\int_0^{(ts_j)^{1/\alpha_j}} \mathbb{1}_{(s_j<R_l^{\alpha_j}/t)}\, r\mathrm{d}r + \int_0^{R_l} r \\ &\qquad \mathbb{1}_{(s_j>R_l^{\alpha_j}/t)}\mathrm{d}r\Bigg] + \int_{R_l}^{(ts_n)^{1/\alpha_n}} \mathbb{1}_{(s_n>R_l^{\alpha_n}/t)}\, r\mathrm{d}r \Bigg\} \\ &= \pi\lambda\Bigg\{ \sum_{j\in\{l,n\}} \mathcal{P}_j \left[t^{2/\alpha_j} \Upsilon_{s_j,2/\alpha_j}(R_l^{\alpha_j}/t) + R_l^2 \bar{F}_{s_j}(R_l^{\alpha_j}/t)\right] \\ &\qquad + t^{2/\alpha_n} \bar{\Upsilon}_{s_n,2/\alpha_n}(R_l^{\alpha_n}/t) - R_l^2 \bar{F}_{s_n}(R_l^{\alpha_n}/t) \Bigg\} \end{aligned} \tag{4.5}$$

where $\Upsilon_{s,n}(x) = \int_0^x s^n f_s(s)\mathrm{d}s$ and $\bar{\Upsilon}_{s,n}(x) = \int_x^\infty s^n f_s(s)\mathrm{d}s$ denote the lower and upper truncated n^{th} moment of s,respectively. $\bar{F}_s(x) = \int_x^\infty f_s(s)\mathrm{d}s$ represents the counterpart cumulative density function (CCDF) of s at x. Recall that $s \sim \ln\mathcal{N}(m,\sigma)$ with mean $m = -0.1\beta\ln 10$ and variance $\sigma = 0.1\vartheta\ln 10$, $\Upsilon_{s,n}(x)$ and $\bar{F}_s(x)$ are given by

$$\begin{aligned}\Upsilon_{s,n}(x) &= \exp(\sigma^2 n^2/2 + nm)\, Q(\frac{\sigma^2 n + m - \ln x}{\sigma}), \\ \bar{\Upsilon}_{s,n}(x) &= \exp(\sigma^2 n^2/2 + nm)\, Q(-\frac{\sigma^2 n + m - \ln x}{\sigma}), \\ \bar{F}_s(x) &= Q(\frac{\ln x - m}{\sigma}). \end{aligned} \tag{4.6}$$

$$\Lambda_a([0,t)) = \pi\lambda \Bigg\{ t^{2/\alpha_n} \exp\left(2(\frac{\sigma_n}{\alpha_n})^2 + \frac{2m_n}{\alpha_n}\right)\left(1 - \mathcal{P}_l\, Q\left(\frac{\sigma_n^2 + m_n - \ln\left(\frac{R_l^{\alpha_n}}{t}\right)}{\sigma_n}\right)\right)$$
$$+ \mathcal{P}_l t^{2/\alpha_l} \exp\left(2(\frac{\sigma_l}{\alpha_l})^2 + \frac{2m_l}{\alpha_l}\right) Q\left(\frac{\sigma_l^2 + m_l - \ln\left(\frac{R_l^{\alpha_l}}{t}\right)}{\sigma_l}\right)$$
$$- \mathcal{P}_l \left[R_l^2 Q\left(\frac{\ln\left(\frac{R_l^{\alpha_n}}{t}\right) - m_n}{\sigma_n}\right) + R_l^2 Q\left(\frac{\ln\left(\frac{R_l^{\alpha_l}}{t}\right) - m_l}{\sigma_l}\right)\right] \Bigg\} \tag{4.7}$$

By substituting (4.6) into (4.4), lemma 4.1 can be obtained. ∎

Therefore, the coverage probability for both the access and backhaul links are given by

$$S_a(\beta) = \mathbb{P}\left[SNR_a > \beta\right] = \mathbb{P}\left[\frac{P_t \psi_a \ell^{-1}(x_a^*)}{\sigma_n^2} > \beta\right]$$
$$= 1 - \exp\left(-\Lambda_a\left(\left[0, \frac{P_t\psi_a}{\sigma_n^2\beta}\right)\right)\right)$$
$$S_b(\beta) = \mathbb{P}\left[SNR_b > \beta\right] = 1 - \exp\left(-\Lambda_b\left(\left[0, \frac{P_t\psi_b}{\sigma_n^2\beta}\right)\right)\right)$$
$$= 1 - \exp\left(-\zeta b_j F_b\left(\beta\right)\right) \tag{4.8}$$

Note that $\Lambda_b\left([0,t)\right)$ is obtained following a similar methodology as in Lemma 1 but by considering the backhaul path loss mapping function, where the mapped point process is $\mathbf{\Psi}_j^c$ instead of $\mathbf{\Psi}$. Hence, the intensity measure $\Lambda_b\left([0,t)\right)$ is similar to (4.7) when replacing λ with $\zeta b_j \lambda$ and the access link parameters with the backhaul link parameters.

4.3.2 Cache Size & Caching Probabilities

Optimal caching scheme

The joint optimal caching distribution b_j^*, $\forall j \in \mathbf{J}$ and cache-enabled fraction ζ^* that maximizes the hit probability can be obtained via solving the following formulation:

$$\max_{\mathbf{b},\zeta} \; S_a(\beta) \sum_{j=1}^{J} a_j \left[\zeta b_j + (1-\zeta b_j)\left(1 - e^{-\zeta b_j F_b(\beta)}\right)\right] \tag{4.9a}$$

$$\text{s.t.} \quad 0 \le b_j \le 1, \quad j = 1, 2, \ldots, J \tag{4.9b}$$

$$\sum_{j=1}^{J} b_j = M = \left\lfloor \frac{M_o}{\zeta^\epsilon} \right\rfloor \tag{4.9c}$$

$$\left(\frac{M_o}{J}\right)^{1/\epsilon} \le \zeta \le 1 \tag{4.9d}$$

The condition in (4.9c) leads to a discrete optimization problem. To guarantee an integer cache size M, the cache-enabled fraction $\zeta \in [\left(\frac{M_o}{J}\right)^{1/\epsilon}, 1]$ is chosen such that $\frac{M_o}{\zeta^\epsilon} \in \mathbb{Z}^+$. We propose solving this problem according to the following iterative algorithm.

i) Initialize ζ within the range $\left(\frac{M_o}{J}\right)^{1/\epsilon} \le \zeta \le 1$.

ii) Given ζ, find the optimal caching distribution b_j^*, $\forall j \in \mathbf{J}$ that maximizes the hit probability.

iii) For the optimal caching distribution b_j^*, find the optimal fraction ζ^*, within

the range $\left(\frac{M_o}{J}\right)^{1/\epsilon} \leq \zeta \leq 1$, that maximized the hit probability.

iv) Iterate between points ii) and iii) until convergence.

For known ζ, the optimization problem in (4.9) turns to the following formulation:

$$\max_{\mathbf{b}} \quad S_a(\beta) \sum_{j=1}^{J} a_j \left[1 - (1 - \zeta b_j)\, e^{-\zeta b_j F_b(\beta)}\right] \tag{4.10a}$$

$$\text{s.t.} \quad 0 \leq b_j \leq 1, \quad j = 1, 2, \ldots, J \tag{4.10b}$$

$$\sum_{j=1}^{J} b_j = \left\lfloor \frac{M_o}{\zeta^{\epsilon}} \right\rfloor = M \tag{4.10c}$$

It is easy to show the concavity of the objective function $\mathcal{H}$ by confirming that the first derivative is ≥ 0 and the second derivative is ≤ 0. Also, the constraints are linear, which imply that the necessity and sufficiency conditions for optimality exist. The KKT Lagrangian function of this problem is given by:

$$\begin{aligned} L(\mathbf{b}, \mathbf{w}, \mu, v) = S_a \sum_{j=1}^{J} a_j \left[1 - (1 - \zeta b_j)\, e^{-\zeta b_j F_b}\right] - \sum_{j=1}^{J} \mu_j\, b_j \\ + \sum_{j=1}^{J} w_j (b_j - 1) + v(M - \sum_{j=1}^{J} b_j). \end{aligned} \tag{4.11}$$

For brevity, we omit the details of finding the optimal caching distribution b_j^*.

Thus, b_j^* at a given cache-enabled fraction ζ is given by

$$b_j^* = \begin{cases} 0 & , v^* < a_j S_a \zeta e^{-\zeta F_b} \left(1 + F_b \left(1 - \zeta\right)\right) \\ 1 & , v^* > a_j S_a \zeta (1 + F_b) \\ \Omega(v^*) & , \text{otherwise} \end{cases} , \tag{4.12}$$

where $\Omega(v^*)$ is the solution of $v^* = a_j S_a \zeta e^{-\zeta b_j^* F_b} \left[1 + F_b \left(1 - \zeta b_j^*\right)\right]$ that satisfies $\sum_{j=1}^{J} b_j^* = \left\lfloor \frac{M_o}{\zeta^\epsilon} \right\rfloor$.

Note that the optimal caching b_j^* can be obtained using bisection method in similar to [1, Algorithm].

Most popular files caching scheme

Another widely accepted and simpler file placement strategy is the most popular file caching (MPC) scheme. Given the descending popularity order of the files indices, the MPC scheme stores all files with indices $j \leq M$ with probability one. Hence, files with indices $j > M$ are never cached at the network edge. The caching probabilities for the MPC scheme are given by

$$b_j = \begin{cases} 1, & j = 1 : \lfloor \frac{M_o}{\zeta^\epsilon} \rfloor \\ 0, & j = \lfloor \frac{M_o}{\zeta^\epsilon} \rfloor + 1 : J \end{cases} \tag{4.13}$$

Therefore the optimal cache size for the MPC can be obtained via the following one parameter optimization problem

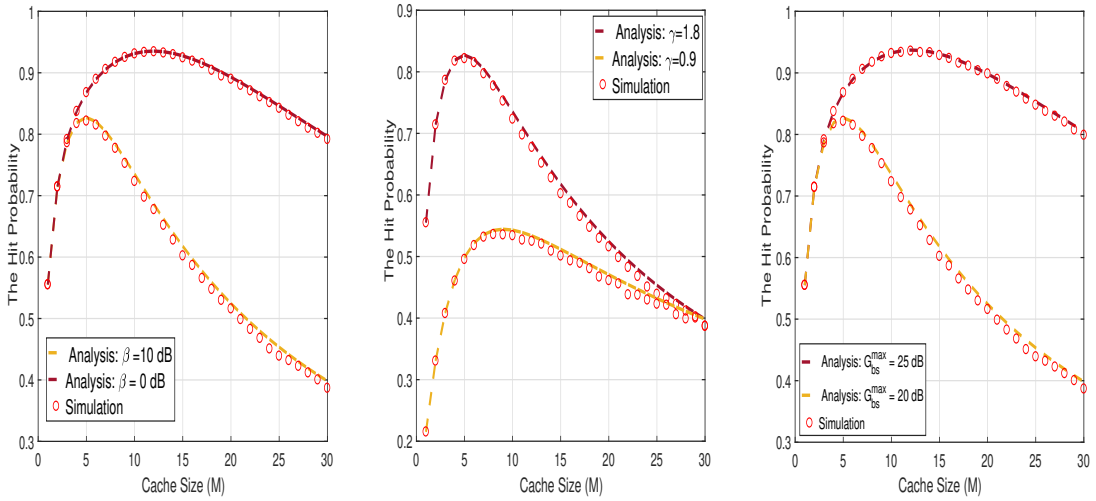
$$\max_{\zeta} \; S_a(\beta) \sum_{j=1}^{\lfloor \frac{M_o}{\zeta^\epsilon} \rfloor} a_j \left[\zeta + (1-\zeta) \left(1 - e^{-\zeta F_b(\beta)} \right) \right] \tag{4.14a}$$

$$\text{s.t.} \quad \left(\frac{M_o}{J} \right)^{1/\epsilon} \leq \zeta \leq 1 \tag{4.14b}$$

which can be solved numerically via looking through all feasible values of ζ.

4.4 Numerical Results

This section validates the developed mathematical model via Monte Carlo simulations. In each simulation run, a PPP of intensity $\lambda = 200$ BSs/Km^2 is generated in a 10×10 km^2 area. Caches with capacities $M = \left\lfloor \frac{M_o}{\zeta^\epsilon} \right\rfloor$ are installed independently at the BSs with probability ζ. CE-BSs independently cache the popular files according to the considered scenarios, namely the optimized caching and MPC schemes. The UEs requests are realized using a Zipf distribution with parameter γ. Each UE is associated to the nearest BS that offers minimum path loss. If the serving BS does not store the requested file by the test UE at the origin, the serving BS utilizes the backhaul to retrieve the requested file from the closest CE-BS that stores it. The simulation is repeated 10^3 times and the hit probability is recored for the test UE. Unless otherwise stated, the mmW network parameters are selected according to [21, Table I] as follows. The mmW frequency and band-



(a) $\gamma = 1.8$ and $G_{bs}^{max} = 20$ dB. (b) $\beta = 10$ dB and $G_{bs}^{max} = 20$ dB. (c) $\gamma = 1.8$, and $\beta = 10$ dB.

Figure 4.3: The hit probability vs. the cache size ($M = \lfloor \frac{M_o}{\zeta} \rfloor$) adopting MPC caching scheme with ($G_u^{max} = 10$ dB, $J = 30$, $M_o = 1$, $\lambda = 200$ BSs/Km^2, and $\epsilon = 1$).

width are set to $f_c = 73$ GHz and $B = 2$ GHz, respectively. The BSs transmit power is $P_t = 30$ dBm, the standard deviation of the path-loss for the access link is $\eta_{\{l,n\}}^a = \{5.2, 7.6\}$, for the backhaul link is $\eta_{\{l,n\}}^b = \{4.2, 7.9\}$, the path loss exponent of the access link is $\alpha_{\{l,n\}}^a = \{2, 3.3\}$, for the backhaul link is $\alpha_{\{l,n\}}^b = \{2, 3.5\}$, where the subscripts l and n denote the LOS and NLOS parameters, respectively. The path loss at reference distance $d_o = 1$m is $\delta_{dB} = 20\log(\frac{4\pi d_o}{\nu}) = 70$ dB, where ν denotes the wavelength. The antennas parameter for the BSs is $\{G_{bs}^{max} = 20$ dB, $G_{bs}^{min} = -2$ dB and $\triangle\phi_{bs} = 10^\circ\}$ and for the UEs is $\{G_u^{max} = 10$ dB, $G_u^{min} = -5$ dB and $\triangle\phi_u = 20^\circ\}$. The single-ball blockage model parameters are $R_l = 200$ m and $\mathcal{P}_l = 0.1$. The caching parameters are chosen as follows; The library size $J = 30$, the smallest cache capacity when $\zeta = 1$, i.e., $M_o = 1$, and the Zipf exponent $\gamma = 1.8$.

Fig. 4.3 shows the hit probability versus the cache size $M = \left\lfloor \frac{M_o}{\zeta^\epsilon} \right\rfloor$ for the MPC caching scheme. It is worth noting that the matching between the simulation and analytical results verifies the proposed mathematical framework. Fig. 4.3 also manifests the cache size/intensity trade-off and indicates that there exists an optimal ζ^* that maximizes the hit probability. In Fig. 4.3a, it can be shown that a higher SNR β requires high intensity of caches, despite their lower capacity of each, due to the low success probability for retrieving files though self-backhauling, and vice versa. Fig. 4.3b shows that decreasing the Zipf exponent requires higher capacities of caches, despite their lower implementation intensities, to satisfy the high number of frequently requested files. Fig. 4.3c illuminates the fact that increasing the BSs antenna gain improves the backhauling efficiency to share files at high SNR thresholds, which allows lower intensity of caches with higher sizes to improve the hit probability.

The optimal caching compared with the MPC is shown in Fig.4.4. The figure both validates the developed mathematical framework and shows that diversifying the file placement though optimal caching leads to higher hit probability. However, the gain of the optimal caching diminishes with large cache size as the effective popular contents are already stored by both schemes. Fig. 4.4 also shows that the optimal cache size (i.e., hence the optimal CE-BSs intensity) can be obtained through a plot over a single parameter over the range $M_o \leq M \leq 1$.

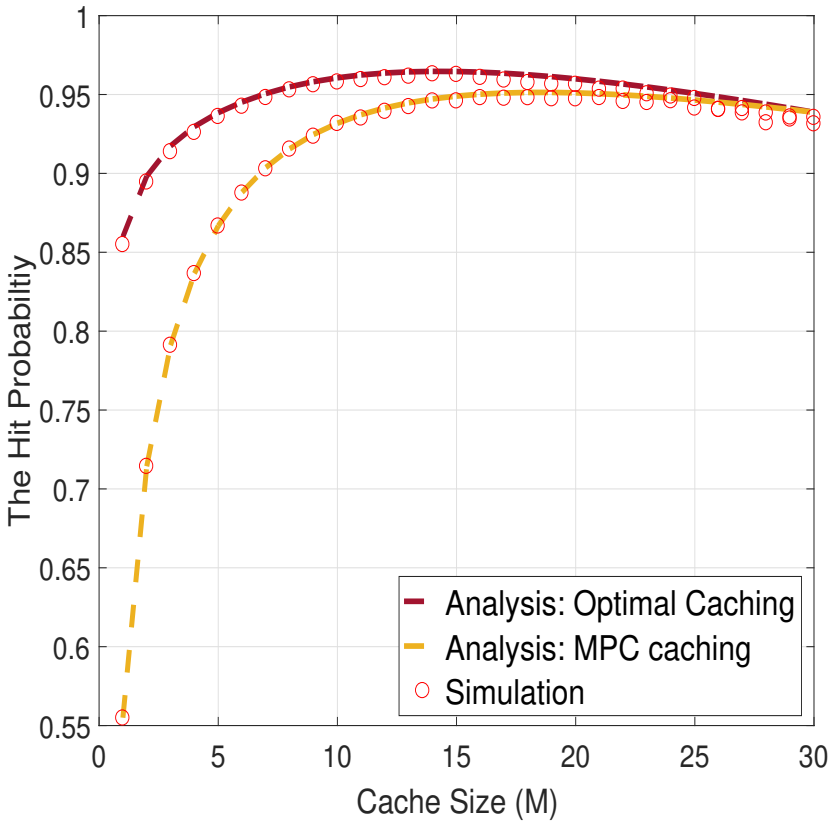


Figure 4.4: The hit probability vs. the cache size for both the MPC and optimal caching schemes with ($J = 30$, $M_o = 1$, $\gamma = 1.8$, $\lambda = 200$ BSs/Km^2, $\beta = 10$ dB, and $\epsilon = 2$).

4.5 Summary

This chapter develops a mathematical framework to study the cache capacity/intensity tradeoff, for a given CAPEX, in an dense self-backhauled mmW fog network. Instead of implementing a cache with a lower capacity at each BS, it may be better to implement larger caches in a fraction of the BSs and share the caches content across BSs though wireless backhaul. To study such tradeoff, the hit probability in the depicted self-backhauled mmW fog network is characterized via stochastic geometry. Assuming an inverse power-law relationship between the

cache size and intensity, an optimization problem is formulated to find the optimal cache intensity and file placement in order to maximize the hit probability. The results show that there exists an optimal balance between the cache size and the intensity of cache enabled BSs, which depends on the network parameters, for the optimized and most popular file caching schemes.

CHAPTER 5

CONCLUSION AND FUTURE WORK

In this chapter, the main contributions of this thesis are summarized, and some future research directions are also presented.

5.1 Summary of Contributions

- We consider an opportunistic spectrum access approach for single-user cache-enabled network. We adopt multi-channel scenario with OSA to improve the network's overall performance. We develop a tractable mathematical model for the hit probability, which is reduced to closed form expression in special cases. Our numerical results both verify the derived analytical paradigms and quantify the gains of the considered multi-channel model. We end up by the achieved improvement of the joint multi-channel cached-enabled network over the single-channel system.

- We develop a mathematical framework, based on stochastic geometry, to characterize the hit probability of a cache-enabled unicast/multicast 5G network with SBS multi-channel capabilities and opportunistic spectrum access (OSA). Integral forms for the hit probability, which are reduced to closed-forms in practical special cases, are obtained. To this end, the sub-optimal caching distribution that maximizes the hit probability is computed. The results manifest the superiority of the multicast transmission mode. The results also signify the importance of OSA and the proposed sub-optimal caching on the hit probability. In particular, at adverse interference conditions, the proposed caching significantly improves the hit probability when compared to the popularity or uniform based caching. When the number of channels increases, the OSA becomes effective and relieves the interference adversity. In this case, the cost of fetching the desired file from a farther SBS decreases and the significance of the caching strategy becomes less distinguishable.

- We propose a flexible caching scheme for the mmWave partially cache-enabled network in which the cache size is inversely proportional to the intensity of cache-enabled SBSs. The optimal caching parameters (i.e., caching distribution and the cache-enabled fraction) that maximizes the hit probability is obtained. To this end, the optimal solution is benchmarked against the most popular caching. The trade-off between the cache-size and the intensity of cache-enabled SBSs is demonstrated via simulation results. The

optimal caching outperforms the most popular caching thanks to the files diversity.

5.2 Proposed Future Work

In above-mentioned work, we study the system performance of the cache-enabled network in terms of the hit probability. Another performance metrics can be considered. One of the suggested future work is modeling *the coverage rate* for the partially cache-enabled network introduced in chapter 4. To characterize the coverage rate, the number of users associated at each SBS is needed. The association model, the self-backhauling, and the dependency on the caching distribution make such analysis a challenge that is needed to be tackled.

In this thesis work, we overlooked the *backhaul phase*, which might be good to be incorporated in the system model. Upon that, network analysis is needed to investigate which of the following scenarios is better and leads to optimistic assessment: *user association to its closest SBS that have the requested file* or *serving the user by its geographically closest SBS by fetching the requested file from the core network via the backhaul link.* This can be characterized in terms of the rate and serving delay.

In the context of mmWave-based network, another direction for future research is to analyze the cache-enabled network in terms of the coverage rate applying spatial division multiplexing (SDM) [86].

Bibliography

[1] Bartlomiej Blaszczyszyn and Anastasios Giovanidis. Optimal geographic caching in cellular networks. In *2015 IEEE International Conference on Communications (ICC)*, pages 3358–3363. IEEE, 2015.

[2] Cisco Visual networking Index. Forecast and methodology, 2016-2021, white paper. *San Jose, CA, USA*, 1, 2016.

[3] Jeffrey G Andrews, Stefano Buzzi, Wan Choi, Stephen V Hanly, Angel Lozano, Anthony CK Soong, and Jianzhong Charlie Zhang. What will 5g be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, 2014.

[4] Jeffrey G Andrews, Tianyang Bai, Mandar Kulkarni, Ahmed Alkhateeb, Abhishek Gupta, and Robert W Heath Jr. Modeling and analyzing millimeter wave cellular systems. *arXiv preprint arXiv:1605.04283*, 2016.

[5] Stephen G Larew, Timothy A Thomas, Mark Cudak, and Amitava Ghosh. Air interface design and ray tracing study for 5g millimeter wave communi-

cations. In *2013 IEEE Globecom Workshops (GC Wkshps)*, pages 117–122. IEEE, 2013.

[6] Zhou Su and Qichao Xu. Content distribution over content centric mobile social networks in 5g. *IEEE Communications Magazine*, 53(6):66–72, 2015.

[7] Chengchao Liang, F Richard Yu, and Xi Zhang. Information-centric network function virtualization over 5g mobile wireless networks. *IEEE network*, 29(3):68–74, 2015.

[8] Kai Fan, Yanhui Ren, Yue Wang, Hui Li, and Yingtang Yang. Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5g. *IET Communications*, 2017.

[9] Min Chen, Yongfeng Qian, Yixue Hao, Yong Li, and Jeungeun Song. Data-driven computing and caching in 5g networks: Architecture and delay analysis. *IEEE Wireless Communications*, 25(1):70–75, 2018.

[10] Karthikeyan Shanmugam, Negin Golrezaei, Alexandros G Dimakis, Andreas F Molisch, and Giuseppe Caire. Femtocaching: Wireless content delivery through distributed caching helpers. *IEEE Transactions on Information Theory*, 59(12):8402–8413, 2013.

[11] Xiaofei Wang, Min Chen, Tarik Taleb, Adlen Ksentini, and Victor Leung. Cache in the air: exploiting content caching and delivery techniques for 5g systems. *IEEE Communications Magazine*, 52(2):131–139, 2014.

[12] Ejder Bastug, Mehdi Bennis, and Mérouane Debbah. Living on the edge: The role of proactive caching in 5G wireless networks. *IEEE Communications Magazine*, 52(8):82–89, 2014.

[13] Negin Golrezaei, Parisa Mansourifard, Andreas F Molisch, and Alexandros G Dimakis. Base-station assisted device-to-device communications for high-throughput wireless video networks. *IEEE Transactions on Wireless Communications*, 13(7):3665–3676, 2014.

[14] Lin Zhang, Ming Xiao, Gang Wu, and Shaoqian Li. Efficient scheduling and power allocation for d2d-assisted wireless caching networks. *IEEE Transactions on Communications*, 64(6):2438–2452, 2016.

[15] Derya Malak, Mazin Al-Shalash, and Jeffrey G Andrews. Optimizing content caching to maximize the density of successful receptions in device-to-device networking. *IEEE Transactions on Communications*, 64(10):4365–4380, 2016.

[16] Zheng Chen, Nikolaos Pappas, and Marios Kountouris. Probabilistic caching in wireless d2d networks: Cache hit optimal versus throughput optimal. *IEEE Communications Letters*, 21(3):584–587, 2017.

[17] Konstantinos Poularakis, George Iosifidis, Vasilis Sourlas, and Leandros Tassiulas. Exploiting caching and multicast for 5G wireless networks. *IEEE Transactions on Wireless Communications*, 15(4):2995–3007, 2016.

[18] Theodore S Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N Wong, Jocelyn K Schulz, Mathew Samimi, and Felix Gutierrez. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE access*, 1:335–349, 2013.

[19] Zhouyue Pi and Farooq Khan. An introduction to millimeter-wave mobile broadband systems. *IEEE communications magazine*, 49(6), 2011.

[20] Amitava Ghosh, Timothy A Thomas, Mark C Cudak, Rapeepat Ratasuk, Prakash Moorut, Frederick W Vook, Theodore S Rappaport, George R MacCartney, Shu Sun, and Shuai Nie. Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks. *IEEE Journal on Selected Areas in Communications*, 32(6):1152–1163, 2014.

[21] Sarabjot Singh, Mandar N Kulkarni, Amitava Ghosh, and Jeffrey G Andrews. Tractable model for rate in self-backhauled millimeter wave cellular networks. *IEEE Journal on Selected Areas in Communications*, 33(10):2196–2211, 2015.

[22] Yongxu Zhu, Gan Zheng, Lifeng Wang, Kai-Kit Wong, and Liqiang Zhao. Content placement in cache-enabled sub-6 ghz and millimeter-wave multi-antenna dense small cell networks. *arXiv preprint arXiv:1801.05756*, 2018.

[23] Yongxu Zhu, Gan Zheng, Kai-Kit Wong, Shi Jin, and Sangarapillai Lambotharan. Performance analysis of cache-enabled millimeter wave small cell networks. *IEEE Transactions on Vehicular Technology*, 2018.

[24] Sudip Biswas, Tong Zhang, Keshav Singh, Satyanarayana Vuppala, and Tharmalingam Ratnarajah. An analysis on caching placement for millimeter/micro wave hybrid networks. *arXiv preprint arXiv:1802.10326*, 2018.

[25] Hesham ElSawy, Ekram Hossain, and Martin Haenggi. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey. *IEEE Communications Surveys & Tutorials*, 15(3):996–1019, 2013.

[26] Hesham ElSawy, Ahmed Sultan-Salem, Mohamed-Slim Alouini, and Moe Z Win. Modeling and analysis of cellular networks using stochastic geometry: A tutorial. *IEEE Communications Surveys & Tutorials*, 2016.

[27] Jeffrey G Andrews, François Baccelli, and Radha Krishna Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Transactions on Communications*, 59(11):3122–3134, 2011.

[28] Anjin Guo and Martin Haenggi. Spatial stochastic models and metrics for the structure of base stations in cellular networks. *IEEE Transactions on Wireless Communications*, 12(11):5800–5812, 2013.

[29] Wei Lu and Marco Di Renzo. Stochastic geometry modeling of cellular networks: Analysis, simulation and experimental validation. In *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 179–188. ACM, 2015.

[30] Negin Golrezaei, Alexandros G Dimakis, Andreas F Molisch, and Giuseppe Caire. Wireless video content delivery through distributed caching and peer-to-peer gossiping. In *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, pages 1177–1180. IEEE, 2011.

[31] Negin Golrezaei, Karthikeyan Shanmugam, Alexandros G Dimakis, Andreas F Molisch, and Giuseppe Caire. Femtocaching: Wireless video content delivery through distributed caching helpers. In *INFOCOM, 2012 Proceedings IEEE*, pages 1107–1115. IEEE, 2012.

[32] Negin Golrezaei, Andreas F Molisch, Alexandros G Dimakis, and Giuseppe Caire. Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution. *IEEE Communications Magazine*, 51(4):142–149, 2013.

[33] Lin Zhang, Zhao Wang, Ming Xiao, Gang Wu, and Shaoqian Li. Centralized caching in two-layer networks: Algorithms and limits. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2016 IEEE 12th International Conference on*, pages 1–5. IEEE, 2016.

[34] Ejder Baştuğ, Mehdi Bennis, Marios Kountouris, and Mérouane Debbah. Cache-enabled small cell networks: Modeling and tradeoffs. *EURASIP Journal on Wireless Communications and Networking*, 2015(1):41, 2015.

[35] Ejder Baştuğ, Marios Kountouris, Mehdi Bennis, and Mérouane Debbah. On the delay of geographical caching methods in two-tiered heterogeneous networks. In *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*, pages 1–5. IEEE, 2016.

[36] Berksan Serbetci and Jasper Goseling. On optimal geographical caching in heterogeneous cellular networks. In *Wireless Communications and Networking Conference (WCNC), 2017 IEEE*, pages 1–6. IEEE, 2017.

[37] Mehrnaz Afshang and Harpreet S Dhillon. Optimal geographic caching in finite wireless networks. In *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*, pages 1–5. IEEE, 2016.

[38] Seong Ho Chae and Wan Choi. Caching placement in stochastic wireless caching helper networks: Channel selection diversity via caching. *IEEE Transactions on Wireless Communications*, 15(10):6626–6637, 2016.

[39] Youjia Chen, Ming Ding, Jun Li, Zihuai Lin, Guoqiang Mao, and Lajos Hanzo. Probabilistic small-cell caching: Performance analysis and optimization. *IEEE Transactions on Vehicular Technology*, 66(5):4341–4354, 2017.

[40] Chenchen Yang, Zhiyong Chen, Yao Yao, and Bin Xia. Performance analysis of wireless heterogeneous networks with pushing and caching. In *Commu-*

nications (ICC), 2015 IEEE International Conference on, pages 2190–2195. IEEE, 2015.

[41] Chenchen Yang, Yao Yao, Zhiyong Chen, and Bin Xia. Analysis on cache-enabled wireless heterogeneous networks. *IEEE Transactions on Wireless Communications*, 15(1):131–145, 2016.

[42] Hesham ElSawy, Ekram Hossain, and Sergio Camorlinga. Multi-channel design for random csma wireless networks: A stochastic geometry approach. In *Communications (ICC), 2013 IEEE International Conference on*, pages 1656–1660. IEEE, 2013.

[43] Hesham ElSawy, Ekram Hossain, and Sergio Camorlinga. Spectrum-efficient multi-channel design for coexisting ieee 802.15. 4 networks: A stochastic geometry approach. *IEEE Transactions on Mobile Computing*, 13(7):1611–1624, 2014.

[44] Hesham ElSawy and Ekram Hossain. Two-tier hetnets with cognitive femtocells: Downlink performance modeling and analysis in a multichannel environment. *IEEE Transactions on Mobile Computing*, 13(3):649–663, 2014.

[45] Ahmed Hamdi Sakr and Ekram Hossain. Cognitive and energy harvesting-based d2d communication in cellular networks: Stochastic geometry modeling and analysis. *IEEE Transactions on Communications*, 63(5):1867–1880, 2015.

[46] Yago Sanchez, Edward Grinshpun, David Faucher, Thomas Schieri, and Sameer Sharma. Low latency dash based streaming over lte. In *Visual Com-*

munications and Image Processing Conference, 2014 IEEE, pages 1–4. IEEE, 2014.

[47] 3rd Generation Partnership Project (3GPP). 2016 [online]. http://www.3gpp.org/specifications/releases/71-release-9.

[48] Jia Guo, Xiangyang Gong, Jie Liang, Shiju Zhang, Mincheng Zhao, Wendong Wang, and Zhiming Li. A hybrid transmission approach for dash over mbms in lte network. In *Global Communications Conference (GLOBECOM), 2015 IEEE*, pages 1–6. IEEE, 2015.

[49] Ying Cui, Dongdong Jiang, and Yueping Wu. Analysis and optimization of caching and multicasting in large-scale cache-enabled wireless networks. *IEEE Transactions on Wireless Communications*, 15(7):5101–5112, 2016.

[50] Ying Cui and Dongdong Jiang. Analysis and optimization of caching and multicasting in large-scale cache-enabled heterogeneous wireless networks. *IEEE Transactions on Wireless Communications*, 16(1):250–264, 2017.

[51] Zitian Wang, Zhehan Cao, Ying Cui, and Yang Yang. Joint and competitive caching designs in large-scale multi-tier wireless multicasting networks. *arXiv preprint arXiv:1706.07903*, 2017.

[52] Syed Tamoor-ul Hassan, Mehdi Bennis, Pedro HJ Nardelli, and Matti Latva-Aho. Caching in wireless small cell networks: A storage-bandwidth tradeoff. *IEEE Communications Letters*, 20(6):1175–1178, 2016.

[53] Derya Malak and Mazin Al-Shalash. Optimal caching for device-to-device content distribution in 5g networks. In *Globecom Workshops (GC Wkshps), 2014*, pages 863–868. IEEE, 2014.

[54] Tianyang Bai and Robert W Heath. Coverage and rate analysis for millimeter-wave cellular networks. *IEEE Transactions on Wireless Communications*, 14(2):1100–1114, 2015.

[55] Marco Di Renzo. Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks. *IEEE Transactions on Wireless Communications*, 14(9):5038–5057, 2015.

[56] Esma Turgut and M Cenk Gursoy. Coverage in heterogeneous downlink millimeter wave cellular networks. *IEEE Transactions on Communications*, 65(10):4463–4477, 2017.

[57] Mustafa Riza Akdeniz, Yuanpeng Liu, Mathew K Samimi, Shu Sun, Sundeep Rangan, Theodore S Rappaport, and Elza Erkip. Millimeter wave channel modeling and cellular capacity evaluation. *IEEE journal on selected areas in communications*, 32(6):1164–1179, 2014.

[58] Xianghao Yu, Jun Zhang, Martin Haenggi, and Khaled B Letaief. Coverage analysis for millimeter wave networks: The impact of directional antenna arrays. *IEEE journal on selected areas in communications*, 35(7):1498–1512, 2017.

[59] Lifeng Wang, Kai-Kit Wong, Shi Jin, Gan Zheng, and Robert W Heath Jr. A new look at physical layer security, caching, and wireless energy harvesting for heterogeneous ultra-dense networks. *arXiv preprint arXiv:1705.09647*, 2017.

[60] Nikolaos Giatsoglou, Konstantinos Ntontin, Elli Kartsakli, Angelos Antonopoulos, and Christos Verikoukis. D2d-aware device caching in mmwave-cellular networks. *IEEE Journal on Selected Areas in Communications*, 35(9):2025–2037, 2017.

[61] Jian Qiao, Yejun He, and Xuemin Sherman Shen. Proactive caching for mobile video streaming in millimeter wave 5g networks. *IEEE Transactions on Wireless Communications*, 15(10):7187–7198, 2016.

[62] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S Kendall, and Joseph Mecke. *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[63] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 1–14. ACM, 2007.

[64] Namyoon Lee, David Morales-Jimenez, Angel Lozano, and Robert W Heath. Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach. *IEEE Transactions on Wireless Communications*, 14(1):230–241, 2015.

[65] Rabe Arshad, Hesham ElSawy, Sameh Sorour, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. Handover management in dense cellular networks: A stochastic geometry approach. In *Communications (ICC), 2016 IEEE International Conference on*, pages 1–7. IEEE, 2016.

[66] Harpreet S Dhillon, Radha Krishna Ganti, François Baccelli, and Jeffrey G Andrews. Coverage and ergodic rate in k-tier downlink heterogeneous cellular networks. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1627–1632. IEEE, 2011.

[67] Tugce Bilen, Berk Canberk, and Kaushik R Chowdhury. Handover management in software-defined ultra-dense 5G networks. *IEEE Network*, 31(4):49–55, 2017.

[68] Gaurav Nigam, Paolo Minero, and Martin Haenggi. Coordinated multipoint joint transmission in heterogeneous networks. *IEEE Transactions on Communications*, 62(11):4134–4146, 2014.

[69] Ahmed Hamdi Sakr and Ekram Hossain. Location-aware cross-tier coordinated multipoint transmission in two-tier cellular networks. *IEEE Transactions on Wireless Communications*, 13(11):6311–6325, 2014.

[70] Hesham ElSawy, Wenhan Dai, Mohamed-Slim Alouini, and Moe Z Win. Base station ordering for emergency call localization in ultra-dense cellular networks. *IEEE Access*, 2017.

[71] Martin Haenggi. *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[72] Tsu T Soong. *Fundamentals of probability and statistics for engineers*. John Wiley & Sons, 2004.

[73] Hesham ElSawy and Ekram Hossain. Channel assignment and opportunistic spectrum access in two-tier cellular networks with cognitive small cells. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 4477–4482. IEEE, 2013.

[74] Mostafa Emara, Hesham ElSawy, Sameh Sorour, Samir Al-Ghadhban, Mohamed-Slim Alouini, and Tareq Y. Al-Naffouri. Stochastic geometry model for multi-channel fog radio access networks. In *Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks (WiOpt), 2017 15th International Symposium on*, pages 1–6. IEEE, 2017.

[75] Rabe Arshad, Hesham ElSawy, Sameh Sorour, Tareq Y Al-Naffouri, and Mohamed-Slim Alouini. Velocity-aware handover management in two-tier cellular networks. *IEEE Transactions on Wireless Communications*, 16(3):1851–1867, 2017.

[76] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products, Seventh Edition*. Academic Press, 2007.

[77] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[78] Mohammed S Elbamby, Cristina Perfecto, Mehdi Bennis, and Klaus Doppler. Edge computing meets millimeter-wave enabled vr: Paving the way to cutting the cord. 2018.

[79] Lifeng Wang, Kai-Kit Wong, Shi Jin, Gan Zheng, and Robert W. Heath Jr. A new look at physical layer security, caching, and wireless energy harvesting for heterogeneous ultra-dense networks. *CoRR*, abs/1705.09647, 2017.

[80] Jeffrey G Andrews, Tianyang Bai, Mandar N Kulkarni, Ahmed Alkhateeb, Abhishek K Gupta, and Robert W Heath. Modeling and analyzing millimeter wave cellular systems. *IEEE Transactions on Communications*, 65(1):403–430, 2017.

[81] Ahmad AlAmmouri, Jeffrey G Andrews, and François Baccelli. Sinr and throughput of dense cellular networks with stretched exponential path loss. *IEEE Transactions on Wireless Communications*, 17(2):1147–1160, 2018.

[82] M. Emara, H. El Sawy, S. Sorour, S. Al-Ghadhban, M. S. Alouini, and T. Y. Al-Naffouri. Optimal caching in 5g networks with opportunistic spectrum access. *IEEE Transactions on Wireless Communications*, pages 1–1, 2018.

[83] Wenqiang Yi, Yuanwei Liu, and Arumugam Nallanathan. Modeling and analysis of mmwave communications in cache-enabled hetnets. *arXiv preprint arXiv:1801.08801*, 2018.

[84] Yongxu Zhu, Gan Zheng, Lifeng Wang, Kai-Kit Wong, and Liqiang Zhao. Content placement in cache-enabled sub-6 ghz and millimeter-wave multi-antenna dense small cell networks. *CoRR*, abs/1801.05756, 2018.

[85] Ming Ding, Peng Wang, David López-Pérez, Guoqiang Mao, and Zihuai Lin. Performance impact of los and nlos transmissions in dense cellular networks. *IEEE Transactions on Wireless Communications*, 15(3):2365–2380, 2016.

[86] Ansuman Adhikary, Ebrahim Al Safadi, Mathew K Samimi, Rui Wang, Giuseppe Caire, Theodore S Rappaport, and Andreas F Molisch. Joint spatial division and multiplexing for mm-wave channels. *IEEE Journal on Selected Areas in Communications*, 32(6):1239–1255, 2014.

Vitae

- Name: Mostafa Mohamed Lotfy Ahmed Emara
- Nationality: Egyptian
- Date of Birth: June, 25th, 1983
- Email: *mostafaemara@kfupm.edu.sa* / *mostafa.emara@f-eng.tanta.edu.eg*
- Permenant Address: Tanta University, Engineering Faculty, Electronics and Electrical Communications Engineering Dept., 31111, Tanta, Egypt.